

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-346362

(43)Date of publication of application : 14.12.1999

(51)Int.Cl.

H04N 7/24
// H04N 5/92

(21)Application number : 11-063139

(71)Applicant : INTERNATL BUSINESS MACH CORP
<IBM>

(22)Date of filing : 10.03.1999

(72)Inventor : NADER MOUSENIAN

(30)Priority

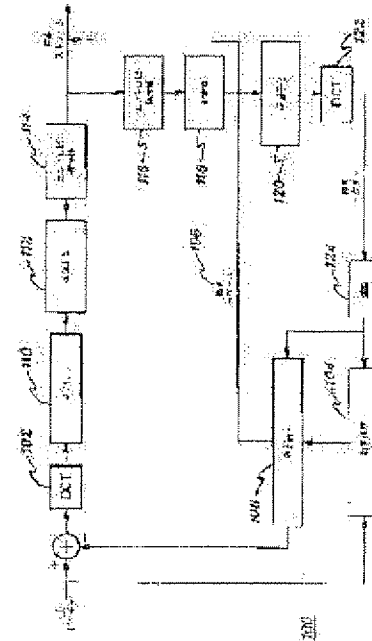
Priority number : 98 44642 Priority date : 19.03.1998 Priority country : US

(54) REAL TIME SINGLE PATH VARIABLE BIT RATE CONTROL METHOD, AND ENCODER

(57)Abstract:

PROBLEM TO BE SOLVED: To attain a real time single path variable bit rate encoding system by encoding at least one frame group of an input stream, and deciding the number of bits to be assigned to another frame group of the input stream through a prescribed processing before encoding the whole input stream.

SOLUTION: The method for processing a frame group comprises a step for encoding at least one frame group of an input stream and a step for deciding the number of bits to be assigned to another frame group of the input stream to be encoded in a variable bit rate mode by using one or more parameter values obtained from the encoding step before encoding the whole input stream. This method can be integrated into a coder for encoding a video stream, and decoding a compressed stream such as an MPEG-2 coder 100.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平11-346362

(43)公開日 平成11年(1999)12月14日

(51) Int.Cl.⁸
H 0 4 N 7/24
// H 0 4 N 5/92

識別記号

F I
H O 4 N 7/13
5/92

$$\begin{array}{c} Z \\ H \end{array}$$

審査請求 有 請求項の数40 O.L (全 28 頁)

(21)出願番号 特願平11-63139

(22)出願日 平成11年(1999)3月10日

(31)優先權主張番号 09/044642

(32)優先日 1998年3月19日

(33) 優先權主張國 美國 (US)

(71)出願人 390009531

インターナショナル・ビジネス・マシーンズ・コーポレーション

INTERNATIONAL BUSINESS
MACHINES CORPORATION

アメリカ合衆国10504、ニューヨーク州
アーモンク (番地なし)

(72)発明者 ネイダー・モウセニアン

アメリカ合衆国13760、ニューヨーク州エ
ンディコット、オデル・アベニュー 322

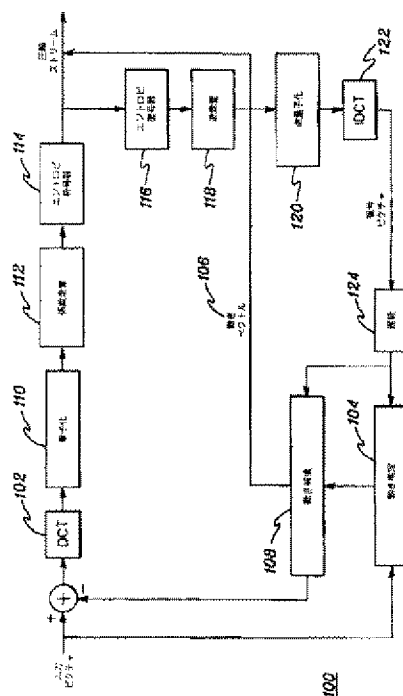
(74) 代理人 弁理士 坂口 博 (外1名)

(54) 【発明の名称】 リアルタイム単一バス可変ビット・レート制御方法及び符号器

(57)【要約】 (修正有)

【課題】 高品質ビデオを提供する、リアルタイム単一パス可変ビット・レート符号化方式の提供。

【解決手段】ビデオ圧縮システムで可変ビット・レート(VBR)MPEG-2符号化を達成する、リアルタイム単一パス可変ビット・レート制御技術が提供される。フレームのシーケンスに対して、ビデオ・インタバル(すなわちGOP)の符号化難度のレベルが、知覚レート-量子化(パーC-パーQ)モデルにより決定される。このモデルは、各ビデオ・インタバルに合成(パーC-パーQ)曲線を割当て、そこからビデオ・インタバルに対するビット数が推定される。この推定は、前に符号化されたビデオ・インタバルから獲得されるパラメータを用いる、因果予測モデルにもとづく。解析中のビデオ・インタバルの総レートから各ピクチャ・タイプのR-Q関係が更新され、ピクチャ・ビットが割当てられる。



【特許請求の範囲】

【請求項 1】フレーム・グループを処理する方法であつて、

入力ストリームの少なくとも 1 つのフレーム・グループを符号化するステップと、

前記入力ストリーム全体を符号化する前に、前記符号化するステップから獲得される 1 つ以上のパラメータ値を用い、可変ビット・レート・モードで符号化される、前記入力ストリームの別のフレーム・グループに割り当てられるビット数を決定するステップとを含む、方法。

【請求項 2】前記別のフレーム・グループを、単一パス可変ビット・レート・モードで、1 つの符号化パスだけにより符号化するステップを含む、請求項 1 記載の方法。

【請求項 3】前記入力ストリームの他のフレーム・グループに対して、前記決定するステップ及び前記別のフレーム・グループを符号化するステップを繰り返すステップを含み、前記他のフレーム・グループの各々に対して決定するステップが、1 つ以上の前に符号化されたフレーム・グループから獲得される 1 つ以上のパラメータ値を使用する、請求項 2 記載の方法。

【請求項 4】前記別のフレーム・グループを符号化するステップが、

前記別のフレーム・グループの 1 つ以上のフレームに割*

$$\bar{C}_k = \left(\frac{R_{k, \text{tot}} N_k}{P_k Z_k} \right) \left(\frac{\xi_k a_1 + \eta_k a_2}{\xi_k + a_2 \left(\frac{R_{k, \text{tot}} N_k}{P_k Z_k} \right)} \right)$$

で与えられ、ここで $R_{k, \text{tot}}$ はビット・バジェット内の残りのビット数、 N_k は入力ストリーム内のフレーム・グループの調整された数、 P_k は符号化される前記別のフレーム・グループの残りのフレームの数、 Z_k は符号化難度の総和をそれぞれ表し、 $\xi_k = \Delta$ (バー C_k) / Δ (バー Q_k) であり、バー C_k は平均ビット、バー Q_k は平均量子化因子を表し、 $\eta_k = \text{バー } C_{k-1, \text{actual}} + \xi_{k-1}$ (バー $Q_{k-1, \text{actual}}$) であり、 a_1 及び a_2 が定数である、請求項 6 記載の方法。

【請求項 9】前記決定するステップが、前記別のフレーム・グループに対して、所定の知覚関数の変換因子を決定するステップと、前記変換因子をビット・レート方程式内で使用し、前記ビット数を決定するステップとを含む、請求項 1 記載の方法。

【請求項 10】前記変換因子を決定するステップが、

【数 3】

* 当てられる理想ビット数を決定するステップと、

復号器バッファにおいて、オーバーフローまたはアンダーフローが生じないように、前記理想ビット数を調整するステップとを含む、請求項 2 記載の方法。

【請求項 5】前記調整するステップが、前記 1 つ以上のフレームに割り当てられる前記理想ビット数を制限するために使用される、1 つ以上のピクチャ限界を決定するステップを含む、請求項 4 記載の方法。

【請求項 6】前記決定するステップが、

10 前記別のフレーム・グループに対して、所定の知覚関数の傾きを調節するステップと、

調節された傾きをビット・レート方程式内で用い、前記ビット数を決定するステップとを含む、請求項 1 記載の方法。

【請求項 7】前記所定の知覚関数が、

【数 1】

$$\bar{C}_k = f^{-1} K F(\bar{Q}_k)$$

で与えられ、ここで $F(\text{バー } Q_k) = a_1 + a_2$ (バー Q_k , β) であり、 a_1 、 a_2 、 β は定数、バー Q_k は量子化因子、及び K は傾きの値をそれぞれ表す、請求項 6 記載の方法。

【請求項 8】前記ビット・レート方程式が、

【数 2】

$$a_l^k = \frac{P_k^{-1} R_{k, \text{tot}}}{\bar{C}_{\text{gop}}} - a_2 \frac{Y_k}{N_k}$$

を使用し、ここで P_k は符号化される前記別のフレーム・グループの残りのフレームの数、 $R_{k, \text{tot}}$ はビット・バジェット内の残りのビット数、バー C_{gop} は前記入力ストリームのレート、 Y_k は 1 フレーム・グループ当たりの平均量子化因子の総和、 N_k は入力ストリーム内のフレーム・グループの調整された数をそれぞれ表し、 a_1 は定数である、請求項 9 記載の方法。

【請求項 11】前記符号化するステップが、前記入力ストリームの第 1 のフレーム・グループ及び第 2 のフレーム・グループを符号化するステップを含み、前記 1 つ以上のパラメータ値が、前記第 1 のフレーム・グループに対する第 1 のビット・レート及び第 1 の量子化因子と、前記第 2 のフレーム・グループに対する第 2 のビット・レート及び第 2 の量子化因子とを含む、請求項 1 記載の方法。

【請求項 1 2】前記決定するステップが、前記第 1 のビット・レート、前記第 2 のビット・レート、前記第 1 の量子化因子、及び前記第 2 の量子化因子を用いる所定の関数により、割当てられる前記ビット数を決定するステップを含む、請求項 1 記載の方法。

【請求項 1 3】前記所定の関数が、

【数 4】

$$\overline{C}_k = \left(\frac{R_{k, \text{tot}} N_k}{P_k Z_k} \right) \left(\frac{\xi_k a_1 + \eta_k a_2}{\xi_k + a_2 \left(\frac{R_{k, \text{tot}} N_k}{P_k Z_k} \right)} \right)$$

で与えられ、ここで $R_{k, \text{tot}}$ はビット・パケット内の残りのビット数、 N_k は入力ストリーム内のフレーム・グループの調整された数、 P_k は符号化される前記別のフレーム・グループの残りのフレームの数、 Z_k は符号化難度の総和をそれぞれ表し、 $\xi_k = \Delta$ (バー C_k) / Δ (バー Q_k) であり、バー C_k は平均ビット、バー Q_k は平均量子化因子を表し、 $\eta_k = \text{バー } C_{k-1, \text{actual}} + \xi_k$ (バー $Q_{k-1, \text{actual}}$) であり、 a_1 及び a_2 が定数である、請求項 1 2 記載の方法。

【請求項 1 4】フレーム・グループを処理する方法であって、

1 つ以上の前に符号化されたフレーム・グループから、複数のパラメータ値を決定するステップと、単一パス可変ビット・レート・モードで符号化されるフレーム・グループに対して、前記複数のパラメータ値の 1 つ以上を用い、所定の関数の傾きを計算するステップ*

$$\overline{C}_k = \left(\frac{R_{k, \text{tot}} N_k}{P_k Z_k} \right) \left(\frac{\xi_k a_1 + \eta_k a_2}{\xi_k + a_2 \left(\frac{R_{k, \text{tot}} N_k}{P_k Z_k} \right)} \right)$$

で与えられ、ここで $R_{k, \text{tot}}$ はビット・パケット内の残りのビット数、 N_k は入力ストリーム内のフレーム・グループの調整された数、 P_k は符号化される前記別のフレーム・グループの残りのフレームの数、 Z_k は符号化難度の総和をそれぞれ表し、 $\xi_k = \Delta$ (バー C_k) / Δ (バー Q_k) であり、バー C_k は平均ビット、バー Q_k は平均量子化因子を表し、 $\eta_k = \text{バー } C_{k-1, \text{actual}} + \xi_k$ (バー $Q_{k-1, \text{actual}}$) であり、 a_1 及び a_2 が定数である、請求項 1 8 記載の方法。

【請求項 2 0】前記所定の関数が、前記別のフレーム・グループの瞬時レートー量子化動作、及び前記別のフレーム・グループの複雑度を示す知覚モデルを表す、請求項 1 8 記載の方法。

【請求項 2 1】前記レートー量子化動作が、

【数 6】

$$\overline{C}_k = -\xi_k \overline{Q}_k + \eta_k$$

により定義され、ここでバー C_k は平均ビット、バー Q_k

*と、

計算された傾き及び前記 1 つ以上のパラメータ値を用い、符号化される前記フレーム・グループのビット・レートを獲得するステップとを含む、方法。

【請求項 1 5】前記可変ビット・レート・モードで符号化される前記フレーム・グループを、当該グループに対して獲得された前記ビット・レートで符号化するステップを含む、請求項 1 4 記載の方法。

【請求項 1 6】前記入力ストリームの他のフレーム・グループに対して、前記決定するステップ、前記計算するステップ、前記獲得するステップ、及び前記符号化するステップを繰り返すステップを含む、請求項 1 5 記載の方法。

【請求項 1 7】前記傾きを計算するステップが、次の関数

$$K_k = f(R_{k, \text{tot}} N_k / P_k Z_k)$$

を使用し、ここで f は前記入力ストリームのフレーム・レート、 $R_{k, \text{tot}}$ はビット・パケット内の残りの数、 N_k は入力ストリーム内のフレーム・グループの調整された数、 P_k は符号化される前記別のフレーム・グループの残りのフレームの数、 Z_k は符号化難度の総和をそれぞれ表す、請求項 1 4 記載の方法。

【請求項 1 8】前記獲得するステップが、所定の関数を用い、前記ビット・レートを獲得するステップを含む、請求項 1 4 記載の方法。

【請求項 1 9】前記所定の関数が、

【数 5】

は平均量子化スケール、 $\xi_k = \Delta$ (バー C_k) / Δ (バー Q_k) であり、 Δ (バー C_k) = バー $C_{k-1, \text{actual}}$ - バー $C_{k-2, \text{actual}}$ 、 Δ (バー Q_k) = バー $Q_{k-2, \text{actual}}$ - バー $Q_{k-1, \text{actual}}$ であり、 $\eta_k = \text{バー } C_{k-1, \text{actual}} + \xi_k$ (バー $Q_{k-1, \text{actual}}$) である、請求項 2 0 記載の方法。

【請求項 2 2】前記知覚モデルが、

【数 7】

$$\overline{C}_k = f^{-1} K F(\overline{Q}_k)$$

により定義され、ここで F (バー Q_k) = $a_1 + a_2$ (バー Q_k) ^{β} であり、 a_1 、 a_2 、 β は定数、バー Q_k は量子化因子、及び K は傾きの値をそれぞれ表す、請求項 2 0 記載の方法。

【請求項 2 3】フレーム・グループを処理する方法であって、

入力ストリームの 1 つのフレーム・グループに割当てられるビット数を、前記フレーム・グループの瞬時レート

一量子化動作、及び前記フレーム・グループの複雑度を示す知覚モデルを用い、決定するステップと、前記フレーム・グループを可変ビット・レート・モードにより、前記決定するステップにより定義されたレートで符号化するステップとを含む、方法。

【請求項 2 4】単一パス可変ビット・レート符号器であって、入力ストリームの少なくとも 1 つのフレーム・グループを符号化するように適応化された符号化エンジンと、前記入力ストリーム全体を符号化する前に、前記少なく

10

とも 1 つのフレーム・グループの符号化から獲得された 1 つ以上のパラメータ値を用い、可変ビット・レート・モードで符号化される、前記入力ストリームの別のフレーム・グループに割り当てられるビット数を決定する手段とを含む、符号器。

【請求項 2 5】前記別のフレーム・グループを符号化するために前記符号化エンジンが、前記別のフレーム・グループの 1 つ以上のフレームに割

20

当てられる理想ビット数を決定する手段と、復号器バッファにおいて、オーバーフローまたはアンダーフローが生じないように、前記理想ビット数を調整する手段とを含む、請求項 2 4 記載の符号器。

【請求項 2 6】前記調整する手段が、前記 1 つ以上のフレームに割り当てられる前記理想ビット数を制限するために使用される、1 つ以上のピクチャ限界を決定する手段を含む、請求項 2 5 記載の符号器。

【請求項 2 7】前記決定する手段が、前記別のフレーム・グループに対して、所定の知覚関数の傾きを調節する手段と、調節された傾きをビット・レート方程式内で用い、前記

30

ビット数を決定する手段とを含む、請求項 2 4 記載の符号器。

【請求項 2 8】前記決定する手段が、前記別のフレーム・グループに対して、所定の知覚関数の変換因子を決定する手段と、前記変換因子をビット・レート方程式内で用い、前記

ビット数を決定する手段とを含む、請求項 2 4 記載の符号器。

【請求項 2 9】前記少なくとも 1 つのフレーム・グループが、前記入力ストリームの第 1 のフレーム・グループ

40

及び第 2 のフレーム・グループを含み、前記 1 つ以上のパラメータ値が、前記第 1 のフレーム・グループに対する、第 1 のビット・レート及び第 1 の量子化因子と、前記第 2 のフレーム・グループに対する、第 2 のビット・レート及び第 2 の量子化因子とを含む、請求項 2 4 記載の符号器。

【請求項 3 0】前記決定する手段が、前記第 1 のビット・レート、前記第 2 のビット・レート、前記第 1 の量子化因子、及び前記第 2 の量子化因子を用いる所定の関数により、割り当てられる前記ビット数を決定する手段を

50

含む、請求項 2 9 記載の符号器。

【請求項 3 1】単一パス可変ビット・レート符号器であって、

1 つ以上の前に符号化されたフレーム・グループから、複数のパラメータ値を決定する手段と、単一パス可変ビット・レート・モードで符号化されるフレーム・グループに対して、前記複数のパラメータ値の 1 つ以上を用い、所定の関数の傾きを計算する手段と、計算された傾き及び前記 1 つ以上のパラメータ値を用い、符号化される前記フレーム・グループのビット・レートを獲得する手段とを含む、符号器。

【請求項 3 2】前記可変ビット・レート・モードで符号化される前記フレーム・グループを、当該グループに対して獲得された前記ビット・レートで符号化する手段を含む、請求項 3 1 記載の符号器。

【請求項 3 3】前記獲得する手段が、所定の関数を用い、前記ビット・レートを獲得する手段を含む、請求項 3 1 記載の符号器。

【請求項 3 4】単一パス可変ビット・レート符号器であって、

入力ストリームの 1 つのフレーム・グループに割り当てられるビット数を、前記フレーム・グループの瞬時レート一量子化動作、及び前記フレーム・グループの複雑度を示す知覚モデルを用い、決定する手段と、前記フレーム・グループを可変ビット・レート・モードにより、前記決定する手段により定義されたレートで符号化するように適応化された符号化エンジンとを含む、符号器。

【請求項 3 5】フレーム・グループの処理をコンピュータに指示するコンピュータ読出し可能プログラム・コード手段を有する、少なくとも 1 つのコンピュータ使用可能媒体を含む装置であって、前記コンピュータ読出し可能プログラム・コード手段が、

入力ストリームの少なくとも 1 つのフレーム・グループを符号化するように、前記コンピュータに指示する手段と、前記入力ストリーム全体を符号化する前に、前記符号化から獲得された 1 つ以上のパラメータ値を用い、可変ビット・レート・モードで符号化される、前記入力ストリームの別のフレーム・グループに割り当てられるビット数を決定するように、前記コンピュータに指示する手段とを含む、装置。

【請求項 3 6】前記別のフレーム・グループを単一パス可変ビット・レート・モードで 1 つの符号化パスだけにより符号化するように、前記コンピュータに指示する手段を含む、請求項 3 5 記載の装置。

【請求項 3 7】前記ビット数を決定するように前記コンピュータに指示する手段が、

前記別のフレーム・グループに対して、所定の知覚関数の傾きを調節するように、前記コンピュータに指示する

手段と、

調節された傾きをビット・レート方程式内で用い、前記ビット数を決定するように、前記コンピュータに指示する手段とを含む、請求項35記載の装置。

【請求項38】前記ビット数を決定するように前記コンピュータに指示する手段が、

前記別のフレーム・グループに対して、所定の知覚関数の変換因子を決定するように、前記コンピュータに指示する手段と、

前記変換因子をビット・レート方程式内で用い、前記ビット数を決定するように、前記コンピュータに指示する手段とを含む、請求項35記載の装置。

【請求項39】前記ビット数を決定するように前記コンピュータに指示する手段が、所定の関数を用い、前記ビット数を決定するように、前記コンピュータに指示する手段を含む、請求項35記載の装置。

【請求項40】フレーム・グループの処理をコンピュータに指示するコンピュータ読出し可能プログラム・コード手段を有する、少なくとも1つのコンピュータ使用可能媒体を含む装置であって、前記コンピュータ読出し可能プログラム・コード手段が、

人力ストリームの1つのフレーム・グループに割り当てられるビット数を、前記フレーム・グループの瞬時レート・量子化動作、及び前記フレーム・グループの複雑度を示す知覚モデルを用い、前記コンピュータに決定するように指示する手段と、

前記フレーム・グループを可変ビット・レート・モードにより、前記手段により定義されたレートで符号化するように、前記コンピュータに指示する手段とを含む、装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は一般に、デジタル・ビデオ圧縮の分野に関して、特に、MPEG-2国際規格を使用するビデオ圧縮システムにおける、リアルタイム単一パス可変ビット・レート符号化にとって好適なレート制御方式に関する。

【0002】

【従来の技術】放送テレビジョン、通信ネットワーク、大衆消費電子製品、及びマルチメディア・コンピュータなどの業界における、様々なデジタル・ビデオ技術の出現は、凄まじい速さで増加している。アプリケーションのこの広範に渡る利用は、デジタル情報の信号処理、編集及びデータ転送が、アナログ表現により同一の作業を実行することに比較して、多大に容易である事実によるものである。しかしながら、より重要なことは、デジタル・ビデオがその大衆性を、ここ数年の間にデジタル・ビデオ圧縮のために作成された幾つかの規格に負っているということである。

【0003】デジタル・ビデオ圧縮ソリューションは、恐

らく、デジタル・ビデオ・プラットフォームの最も重要な要素である。デジタル・ビデオは非圧縮形式では、膨大な量の情報を含むことが知られているので、その処理、記憶及び伝送は、不可能でないにしても、非常に時間を費やし、高価となる。その結果、内容の知覚的品質を保ちながら、圧倒的な量のデータを低減するために、デジタル・ビデオ圧縮技術が必要とされる。次に、データを再生のために圧縮解除するために、互換のビデオ伸張方式が採用される。

【0004】Moving Pictures and Expert Groupにより形成されたMPEG-2国際規格(ISO/IEC 13818-2, "Information technology - Generic coding of moving pictures and associated audio information: Video, 1996")に記載)は、前述の業界の圧縮の必要性を満足するために意図されたものである。この規格は、符号化ストリームの構文及び復号技術を指定する。これは符号器の設計者が、自分自身の圧縮パラメータのセットを導出することを可能にする大きな柔軟性を提供し、従って、彼ら自身の製品を他の製造業者により開発された残りの符号器から区別することを可能にする。

【0005】MPEG-2符号化システムの効率は、固定の通信帯域幅を介して伝送される、またはデジタル媒体に記憶される知覚ビデオの忠実度により評価される。デジタル衛星システムなどの一部のアプリケーションでは、複数のプログラムが単一の大きなストリームに多重化されて同報され、MPEG-2符号器のバンクが、全てのプログラムをモニタ及び符号化するために使用される一方で、全ての受信チャネルの品質を維持する。MPEG-2ストリームは、非同期転送モード(ATM)・ネットワークで見られるように、固定の通信帯域幅または動的な帯域幅を介して伝送される。記憶媒体のサイズが限られる家庭用ビデオまたはPCアプリケーションでは、ユーザができるだけ多くの情報を蓄積することが好ましく、効率的な符号器がこれらの場合では特に有用である。

【0006】ほとんどのMPEG-2符号器は、一定ビット・レート(CBR)・モードで動作するように開発され、ビデオ・ストリームの平均レートは、終始ほぼ同一である。ビデオ・ストリームは、様々なタイプの複数のピクチャまたはフレームを含み、それらにはIピクチャ・タイプ、Bピクチャ・タイプ及びPピクチャ・タイプが含まれる。ピクチャはそのタイプに応じて、ビデオ・ストリームの設定ターゲット・レートよりも多くのビットまたは少ないビットを消費し得る。CBRレート制御方法は、所望の平均ビット・レートが満足され、高品質のビデオ・シーケンスが表示されるように、異なるピクチャ・タイプのストリーム間でビット比率を維持する責任を負う。

【0007】他のMPEG-2符号器などの他の符号器は、可変ビット・レート(VBR)・モードとして定義

される異なるモードで動作する。可変ビット・レート符号化は、各圧縮ピクチャがイントラ・ピクチャ特性及びインタ・ピクチャ特性の複雑度にもとづき、異なる量のビットを有することを可能にする。例えば、単純なピクチャ内容（例えば色テスト・パターンなど）を有するシーンの符号化は、複雑なピクチャ内容（例えば複雑した町並みなど）を有するシーンと比較して、遥かに少ないビットの消費により同一のピクチャ品質を達成できる。

【0008】ビデオを特徴づけるために必要とされる情報の量、及び符号化プロセスを効果的に向上させるための情報を解釈するために必要とされるアルゴリズムの複雑性により、VBR符号化は従来、2つ以上のパスを用いて非リアルタイムに達成された。第1のパスでは、符号化が実行され、統計が収集されて分析され、第2のパスでは、分析結果を用いて符号化プロセスが制御される。これは高品質圧縮ビデオ・ストリームを生成するが、リアルタイム操作を可能にせず、また単一パス符号化を可能にしない。

【0009】

【発明が解決しようとする課題】従って、高品質ビデオを提供する、リアルタイム単一パス可変ビット・レート符号化方式が待望される。更に、ハードウェアにより容易に実現可能な、リアルタイム単一パス可変ビット・レート符号化方式が待望される。

【0010】

【課題を解決するための手段】本発明によれば、フレーム・グループを処理する方法の提供により、従来技術の欠点が克服され、追加の利点が提供される。例えば、本方法は、入力ストリームの少なくとも1つのフレーム・グループを符号化するステップと、入力ストリーム全体を符号化する前に、入力ストリームの別のフレーム・グループに割り当てられるビット数を決定するステップとを含む。別のフレーム・グループは、可変ビット・レート・モードで符号化される。前記決定するステップは、符号化から獲得される1つ以上のパラメータ値を使用する。

【0011】本発明の別の実施例では、本方法は、別のフレーム・グループを単一パス可変ビット・レート・モードで符号化するステップを含み、そこでは別のフレーム・グループが1つの符号化パスだけにより符号化される。

【0012】更に本発明の別の実施例では、前記決定するステップが、別のフレーム・グループに対して、所定の知覚関数の傾きを調節するステップと、調節された傾きをビット・レート方程式内で使用することにより、割り当てられるビット数を決定するステップとを含む。

【0013】更に本発明の別の実施例では、前記決定するステップが、別のフレーム・グループに対して、所定の知覚関数の変換因子を決定するステップと、変換因子をビット・レート方程式内で使用することにより、ビッ

ト数を決定するステップとを含む。

【0014】更に本発明の別の実施例では、前記決定するステップが、入力ストリームの1つ以上の不安定条件を補償するステップを含み、更に他の実施例では、前記決定するステップが、所定の関数を用いてビット数を決定するステップを含む。

【0015】本発明の別の態様では、フレーム・グループを処理する方法が提供される。例えば、この方法は、1つ以上の前に符号化されたフレーム・グループから、複数のパラメータ値を決定するステップと、単一パス可変ビット・レート・モードで符号化されるフレーム・グループに対して、複数のパラメータ値の1つまたは幾つかを用い、所定の関数の傾きを計算するステップと、計算された傾き及び複数のパラメータ値の1つまたは幾つかを用い、符号化されるフレーム・グループに対するビット・レートを獲得するステップとを含む。

【0016】更に本発明の別の態様では、フレーム・グループを処理する方法が提供される。例えば、この方法は、入力ストリームのフレーム・グループに割り当てられるビット数を決定するステップを含む。この決定ステップは、フレーム・グループの瞬時レート・量子化動作、及びフレーム・グループの複雑度を示す知覚モデルを使用する。この方法は更に、可変ビット・レート・モードを用い、決定ステップにより定義されるレートにより、フレーム・グループを符号化するステップを含む。

【0017】更に本発明の別の態様では、単一パス可変ビット・レート符号器が提供される。例えば、この符号器は、入力ストリームの少なくとも1つのフレーム・グループを符号化するように適応化された符号化エンジンと、入力ストリーム全体を符号化する前に、入力ストリームの別のフレーム・グループに割り当てられるビット数を決定する手段とを含む。別のフレーム・グループは、可変ビット・レート・モードで符号化される。前記決定する手段は、符号化エンジンから獲得される1つ以上のパラメータ値を使用する。

【0018】更に本発明の別の態様では、単一パス可変ビット・レート符号器が提供され、これは例えば、1つ以上の前に符号化されたフレーム・グループから、複数のパラメータ値を決定する手段と、単一パス可変ビット・レート・モードで符号化されるフレーム・グループに対して、複数のパラメータ値の1つまたは幾つかを用い、所定の関数の傾きを計算する手段と、計算された傾き及び複数のパラメータ値の1つまたは幾つかを用い、符号化されるフレーム・グループに対するビット・レートを獲得する手段とを含む。

【0019】更に本発明の別の態様では、単一パス可変ビット・レート符号器が提供され、これは例えば、フレーム・グループの瞬時レート・量子化動作、及びフレーム・グループの複雑度を示す知覚モデルを使用し、入力ストリームのフレーム・グループに割り当てられるビット

数を決定する手段と、可変ビット・レート・モードにおいて、決定手段により定義されたレートにより、フレーム・グループを符号化するように適応化された符号化エンジンとを含む。

【0020】本発明は特に、デジタル・ビデオ・ディスク(DVD)、デジタル・ビデオ・カセット・レコーダ(DVCR)、コンピュータ・マルチメディア、及びデジタル・カムコーダ用に家庭で製作する場合のように、ビデオを予め分析または符号化できないアプリケーションにおいて有用である。単一パスVBR符号器は、前に符号化されたピクチャだけを用いて、符号化されるピクチャの複雑度に関して推定を行う。これは高性能DVDマスタリングで使用される、複数パスVBRビデオ圧縮と対照的であり、そこではビデオ・ソースが予め分析され、特定の記憶媒体にとって最適な圧縮ストリームを生成するために、何度も符号化される。従って、単一パスVBR符号器は因果予測モデルに頼ることにより、ストリームの平均レートを動的に調整し、高忠実度のビデオを生成する。

【0021】

【発明の実施の形態】

【数8】

\bar{X}

は、本明細書では \bar{X} と記載する。

【0022】本発明の原理に従えば、リアルタイム単一パス可変ビット・レート(VBR)符号化のためのレート制御方法が提供され、これはハードウェアにより容易に実現可能で、例えばMPEG-2圧縮システムに組み込まれ得る。

【0023】MPEG-2符号器100のブロック図の1例が、図1に示される。特に図1は、ビデオ・ストリームを符号化し、圧縮ストリームを復号するために符号器により使用される、様々なイメージ処理タスク及び信号処理タスクを示す。

【0024】MPEG-2符号器100は高機能な情報プロセッサであり、空間的イメージ分析器及び時間的イメージ分析器、量子化器、エントロピ符号器、及び選択アプリケーションの要求に合う完全準拠のビデオ・ストリームの生成を保証するレート制御ポリシを含む。ビデオ・ストリームの各輝度ピクチャが、16×16画素の非オーバーラップ・ブロックに区分化され、これが更に8×8ブロックに分割される。2つのクロミナンス・ピクチャが、4:2:2または4:2:0クロミナンス・サンプリングに従い、4つのまたは2つの8×8ブロックを生成する。全ての輝度またはクロミナンス・ブロックを含むセットは、マクロブロック(MB)と呼ばれる。4:2:2マクロブロックは8つのブロックを有し、4:2:0マクロブロックは6つのブロックを有する。画素の各8×8ブロックが、2次元離散コサイン変

換(DCT)102を用い、周波数ドメインに変換される。DCTは高優先順位の(低)周波数係数、及び低優先順位の(高)周波数係数のクラスを獲得することにより、入力ブロックに対して、エネルギー圧縮を実行するように意図される。これはイメージ・ブロックの空間的冗長性の一部を除去するための、高速且つ効率的な方法である。

【0025】時間的冗長性は、動き推定(ME)104と呼ばれる技術により除去される。この機構により、前のピクチャに属する一致マクロブロックを探索窓内に移動することにより、任意の輝度マクロブロックが予測される。最良のマクロブロック・マッチを探索するプロセスが、動き推定として定義される。次に、この最良のマッチ(予測)が考慮中のマクロブロックから減算され、DCT変換がマクロブロック差分に適用される。更に、予測されたマクロブロックの位置を指し示す動きベクトル106が生成される。動きベクトル(MV)は、符号器及び復号器の両方により、一致マクロブロックを動き補償し(108)、最良のマッチを突き止めるために使用される。

【0026】動き推定はマクロブロックの輝度部分に対してだけ実行され、結果の動きベクトルがスケールリングされて、クロミナンス・ブロック対を動き補償するために使用される。動きベクトルの符号化は、差分パルス符号変調(DPCM)として知られる技術を用いて実行され、そこでは、動きベクトルが前に符号化されたマクロブロックの動きベクトルを用いて予測され、動きベクトルとその予測との差分だけが符号化される。動きベクトルの差分は、MPEG-2規格のルックアップ・テーブル内で定義されるコード・ワードにより表される。

【0027】MPEG-2規格は、図2に示されるように、I(イントラ符号化)タイプ・ピクチャ、P(予測)タイプ・ピクチャ、及びB(双方向予測)タイプ・ピクチャを使用する。Iタイプは、そのピクチャ内の情報だけを使用するのに対して、Pタイプは、前に符号化されたIタイプまたはPタイプに対して実行される動き推定タスク及び動き補償タスクによりマクロブロック200を予測し、符号化される予測誤差を計算する。これは効率的な圧縮方法である。なぜなら、動き補償されたマクロブロック差分は、原マクロブロックよりも少ない情報を含むからである。

【0028】Pタイプはしばしば、前方予測ピクチャと呼ばれる。なぜなら、それらは前のピクチャを用いて、時間的に前方の予測を行うからである。例えば図2において、動きベクトルMV₁ 202は、ピクチャP内のマクロブロック200に対する最良のマッチを、ピクチャI内の探索窓204(図2で破線により示される)内で見出すために使用される。

【0029】Bタイプ・ピクチャは、前方予測及び後方予測のために、先行及び将来の両方のIタイプまたはP

10

20

30

40

50

タイプからの情報を使用し、特に異なる方向に異なる速度で移動するオブジェクトを推定するために有用である。更に、Bピクチャにおいて使用される双方向動き推定方式は、Pタイプで使用される因果予測に比較して、雑音のある環境においてよりロバストである。これは両方の予測が、それらの公称値を平均することにより、同時に使用されるからである。図2の例では、ピクチャB1内のマクロブロック200に対する動き推定のタスクが、 MV_{B1} により表される前方予測と、 MV_{B2} により表される後方予測とを含む。ピクチャB2では、前方予測及び後方予測が、それぞれ MV_{B1} 及び MV_{B2} により表される。

【0030】様々なピクチャ・タイプが、符号器100により、幾つかのステップにおいて量子化される(図1の110)。最初に、ピクチャ・タイプに固有の量子化マトリックスが 8×8 DCTブロックに適用される。低周波数係数の方を重要視するように、マトリックスの要素(重み)が選択される。これらの係数はより多くの情報を含み、ピクチャの知覚品質により大きく寄与する。次に、量子化スケール・ファクタが計算され、マクロブロックがイメージの局所的なアクティビティ、ピクチャ・タイプの複雑度、及びバッファ充填度にもとづき、異なるクラスに分類される。この指標は、ピクチャに割り当てられるビット数と、その知覚品質とに直接関連付けられる。実際のビット数は、MPEG-2規格において定義されるプロシーダに従い、最初に各DCTブロックを走査し(112)、次に量子化された交流(AC)係数を一緒にグループ化することにより獲得される。各グループは、ハフマン・ルックアップ・テーブルを用いることにより、固有の変長コード(VLC)によりエントロピ符号化される(114)。直流(DC)係数は、DPCM(MVの符号化に類似)及び様量子化を用いて符号化される。

【0031】エントロピ符号器114の出力は、エントロピ復号器116に入力される。復号器116の出力は、逆走査118、逆量子化120、及び逆離散コサイン変換(IDCT)122を介して、差分マクロブロックの有損失バージョンを再構成する。次に、復号されたピクチャが既知のように、遅延124を介して、動き推定104及び動き補償108に渡される。

【0032】圧縮ストリームの瞬時レートは時間と共に変化するので、レートの変動を吸収し、固定チャネル・レートを圧縮ピクチャ・ビットに変換し、それを実際の復号器に引き渡して処理するために復号器バッファが使*

$$R^x(Q^x) = X^x / Q^x$$

【0036】ここで X^x は、各ピクチャ・タイプの複雑度を示す所定のパラメータである。

【0037】しかしながら、ソースの性質は時間と共に変化するので、新たな複雑度パラメータが各ピクチャ・タイプの符号化以前に決定される。このパラメータは通

*用される。従って、本発明のレート制御の1実施例では、後述のように、ピクチャの複雑度及びイメージ・マクロブロックの局所的アクティビティと共に、バッファ占有レベルが追跡される。

【0033】符号器100は更に、選択アプリケーションの要求に合った完全準拠のストリームを生成するために使用される、レート制御ポリシを含む。本発明の原理によれば、符号器100のレート制御ポリシは、入力ストリーム(例えばビデオ・ストリーム)のフレームを符号化するために使用される、可変ビット・レート(VBR)符号化方式を含む。しかしながら、本発明の理解に役立てるために、本発明のVBR符号化方式について述べる前に、幾つかのレート制御原理について述べることにする。これらの原理は、一定ビット・レート(CBR)符号化方式に関連する。

【0034】1. CBRレート制御原理: MPEG-2シーケンスは通常、ピクチャのグループ(GOP)またはフレーム・グループなどの小インタバルに区分化され、これが更にピクチャ・タイプI、P、Bにより類別される。1GOP当たりのビット数は、Iピクチャ割当てが、Pよりも多くなるように分配される。これはPピクチャがその内容を推定するために動き推定技術を使用し、その結果、原ソースよりも低いエントロピを有する、動き補償されたフレーム差分(MCFD)が符号化されるからである。Bピクチャは最も少ない量のビットを使用する。なぜなら、それらの動き推定技術は、Pのそれよりも集約的であるからである。このことは、異なるタイプのピクチャが符号化されるときに、GOP内で同じピクチャ品質を維持するための基本を提供する。更に、Bピクチャのビット割当てでは低減され得る。なぜなら、それらは他のピクチャを推定するために使用されないからである。符号化されたGOPは、Iピクチャ・タイプにより開始する。ピクチャ・タイプx($x \in \{I, P, B\}$)に割り当てられるビットの量は、そのタイプのピクチャ品質において知覚される歪み量と逆の関係にある。

【0035】文献(すなわちJ. L. Mitchell, W. B. Pennebaker, C. E. Fogg及びD. J. LeGallによる"Mpeg Video Compression Standard", Chapman and Hall, New York, 1977)における実験結果は、ソース(またはピクチャもしくはフレーム)のレートRと量子化因子Qとの間には、類似の動作が存在することを示唆しており、これは単純化された形式では、次のようになる。

【数9】

$$1)$$

常、過去の符号化パラメータ、例えばビット、量子化因子、特定の先取り統計などにもとづき計算される。MPEG-2ストリームの各GOPに対して、ビット数Cは次のように設定される。

【数10】

(9)

特開平11-346362

$$\sum_x N^x R^x = C_k$$

 $x = I, P, B$ 16
2)

【0038】ここで指数 k はGOP番号、 x はピクチャ・タイプ、 N^x はGOP内のタイプ x のピクチャの数、 R^x はタイプ x のターゲット・ピクチャ・ビットである。

【0039】CBRシーケンスでは $C_k = C_{\text{seq}}$ であり、ここで C_{seq} は固定GOPビットである。所与の C_k に対して、式2)の条件で量子化スケーラの平均合計

*を最小化することにより、GOPのビデオ品質が最大化される。

【0040】式2)の制約で Ψ を最小化する代わりに、この条件が除去され、ラグランジュ乗数 λ がラグランジュ・コスト T を最小化するために使用される。

【数12】

【数11】

$$\Psi = \frac{\sum_x N^x Q^x}{\sum_x N^x}$$

3)

*

$$T = \Psi - \lambda C_k$$

4)

【0041】式1)で述べられるレート-量子化モデルの支援により、各ピクチャ・タイプのターゲット・ビット※

※トが次のように推論される。

【数13】

$$R^x = \frac{X^x C_k}{\sum_x X^x N^x}$$

5)

【0042】式5)のターゲットは、理想のピクチャ・ビットを表し、実際のビットはほとんどいつもこれから逸脱する。累積誤差が計算され、一定ビット・レート制御方式にフィードバックされ、最終MPEG-2ビット・ストリームが平均ビット・レートまたは総ビット・バジェット(budget)に適合するように保証する。ここで C_k 、ideal及び C_k 、actualが、それぞれGOP k に対する理想の及び実際のビットを表わし、 $\delta_{k,gop} = C_k$ ★

★、actual $- C_k$ 、idealが両者間の差を表すものとする。

更に、 R_k 、ideal及び R_k 、actualが、それぞれピクチャ i の理想の及び実際のビットを表し、 $\delta_{i,pic} = R_k$ 、actual $- R_k$ 、idealが両者間の差を表すものとする。

n 個のピクチャが符号化された後、総累積誤差は次のように計算される。

【数14】

$$\sum_{k=0}^{n_g-1} \delta_{k,gop} + \sum_{i=0}^{n-n_g G-1} \delta_{i,pic} = \Delta_{n-1,gop} + \Delta_{n-1,pic} \quad 6)$$

【0043】GOPのサイズが $G = \sum_x N^x$ により与えられ、 $n_g = \lceil n/G \rceil$ は符号化されたGOPの数である。全ての処理されたGOPに対するサブ誤差累積は、 $\Delta_{n-1,gop}$ により与えられ、 $\Delta_{n-1,pic}$ は符号化順序の最後の、しかしまだ終了されていないGOPに対するサブ誤差累積である。次に、前に符号化されたピクチャから☆

☆生じたビットの生成過剰または生成不足に対して、理想ピクチャ・ターゲットが調整される。GOP ($k = n_g$)に属するピクチャ n を符号化する前の新たな理想ビットは、次のようになる。

【数15】

$$R_{n,ideal}^x = \frac{X_{n-1}^x (C_{gop} - \alpha \Delta_{n-1})}{\sum_x X_{n-1}^x N^x} \quad 7)$$

【0044】ここで C_{gop} はCBR GOPビットの固定数、 α は例えば、1、0、 $\Delta_{n-1} = \Delta_{n-1,gop} + \Delta_{n-1,pic}$ である。 α は、この調整がどれくらい積極的に実行されるかを示す定数である。各ピクチャが符号化された後、複雑度 X^x が更新され、式7)にもとづく新たなターゲットが次のピクチャに対して計算される。このターゲットは後述のように、復号器バッファがオーバー

フロー条件またはアンダーフロー条件に遭遇しないように、特定の制約を満たすべきである。

【0045】最後に、式1)の双曲線上で定義される量子化スケーラ Q が獲得される。各ピクチャ・タイプはそれ自身の合成($R = Q$)曲線を有し、更に、全てのフレームが同様に知覚されるように、各ピクチャ・タイプに対して Q 因子が僅かに調整され得る。マクロブロック・

レベルのレート制御に対して知覚効果が組み込まれる態様については、符号化タスクの間にQ因子の調節により、理想ターゲットが満足される態様同様、ここでは述べられない。しかしながら、これらについては、1993年4月付けのISO/IEC JTC1/SC29/WG11/N0400、"Test Model 5"で述べられている。

【0046】以上、様々なCBR原理について述べてきた。次に、VBR及び特に、本発明のリアルタイム単一パスVBR方式について述べることにする。

【0047】2. VBRビデオ：第1章で述べたCBRレート制御方式は、固定ターゲット C_{rate} に対して、量子化パラメータを調節することにより、GOP内で一定の品質が達成される事実により動機付けされた。時とともにCBRレート制御方式を安定化し、所望のレートを維持するために、任意の統計的变化またはビット派生物が利用される。更に、MPEG-2 CBR符号器は、普遍定数及び所定の初期複雑度のセットを用いて、異なるピクチャ・タイプ間で割当てられるビット量の間の特定の比率を維持する。しかしながら、GOPセグメントまたはビデオ・セグメントなどの連続ピクチャのグループを、複雑度のハードネスまたはソフトネスに関して、異なるタイプの時間インタバルに分類する成果は、リアルタイム単一パスMPEG-2符号化では制限された。高忠実度の結果を生成することが要求されるビデオのハードネス（ソフトネス）は、多量（少量）のビットにより定義される。複数パスCBR符号化または複数パスVBR符号化では、こうした情報は既知であり、従って品質改善が達成され得る。

【0048】しかしながら、本発明では、GOPビットをオン・ザ・フライ式に調整することにより、複雑度レベルがCBR符号化のレート制御方法に追加される。こ*

$$R_s \propto (a_1 + a_2 Q_s \beta) \quad (8)$$

【0051】ここで a_1 は、知覚ライン301がバーC軸302と交わるポイント300（図3）であり（例えば $a_1 = 0.2$ ）、 a_2 は知覚ライン301の傾きであり（例えば $a_2 = 0.6$ ）、 β は知覚モデルが線形か非線形かを示す（例えば $\beta = 1$ は線形を示す）。

【0052】式8)の $R_s \propto Q_s$ 関係は、全てのタイプのビデオ・インタバルが等しく知覚されるべきであるという基準にもとづき推論される。更に、これは大きな量子化スケールQを生成する困難なビデオ・セグメントは、平均ビット・レートの圧縮ストリームよりも多くのビッ*

$$R_s = K (a_1 + a_2 Q_s \beta) = K F(Q_s) \quad (9)$$

【0054】特に、本発明の1方式によれば、定数K（Mビット／秒で測定される）が各ビデオ・インタバル（すなわちGOP）に対して調節され、圧縮ストリームの平均レートが所望のターゲット・レートに合うように保証する。Kを調節する1方式が後述される。

【0055】2. 2 VBRレート制御方式1：単一パスVBR符号器の効率、そのレート制御方式がどのく

*の問題は特に挑戦的である。なぜなら、いかなるビデオ・ソースの統計的性質も事前に知れていないか、時間と共に変化するので、ソース分散の真の推定が不可能になるからである。本発明のリアルタイム単一パスVBR符号器は、（バーC—バーQ）知覚モデルを用い、各々が式1)により定義される特定の双曲線に対応するビデオ・セグメントの、ハードネスまたはソフトネスを区別する。この時、実際の（バーC、バーQ）対が、この双曲線に沿って計算される。

【0049】2. 1 レート—量子化知覚モデル：本発明のVBR方式は、各ビデオ・セグメント（すなわち各GOP）が符号化難度のレベルに関連付けられる事実により、概念的に動機付けされ、この符号化難度は総ピクチャ・ビット、量子化スケール、空間的アクティビティ、時間的アクティビティ、信号対雑音比、またはこれらの組み合わせなどの、様々なソース統計または圧縮パラメータにより測定され得る。より多くのビットが、ハイ・レベルの符号化難度を有するビデオ・セグメントに割当てられるべきである。これは、ソースの複雑度には関係無しに、固定数のビットが各GOPに割当てられるCBRレート制御方式とは異なるアプローチである。研究成果によれば、複雑な、適度な、そして簡単な資料を含む多数のテスト・ケースにおいて、ビデオ・インタバルのレートと量子化スケールとの間には、強い相関が存在することが判明した。この相関については、例えば、1998年1月26日付けの米国特許出願第09/013149号（対応日本国出願は特願平11-6060）で述べられている。

【0050】特に、この相関は次の式により示される。

【数16】

$$(8)$$

※トを消費するべきであり、簡単なセグメントはより少ない数のビットを使用することを示唆する。この関係は、ロバストな単一パスVBR符号化方式を公式化する構成ブロックを提供する。なぜなら、それはビデオ・ソースの可変性を利用するからである。

【0053】前記の関係をj用い、各インタバル（例えばGOP）に割当てられる実際のビット数が、本発明の原理に従い、次の知覚モデルの傾きKにより決定される。

【数17】

$$(9)$$

らい速く、ビデオ・ストリームのソフトネスまたはハードネスを学習し、自身をそれに調整することができるかにより評価される。イメージの不連続性または特殊効果が発生する領域では、ピクチャ品質の劣化が最小化されるべきである。単一パス符号化では、イメージ統計が前に分析及び符号化されたピクチャにより制限されるので、レート制御方式の学習レートは、将来のビデオ・イ

ンタバルの内容を予測するのに十分であるべきであるが、アルゴリズム的な不安定性を生じるほど積極的であってはならない。この2重の問題を解決する1つの方法は、時間インタバルごとに（すなわちGOPごとに）、符号化ストリームの品質を調整し、レート制御方式にその時間インタバル内の各ピクチャのローカル内容を学習させることである。

【0056】本発明の原理に従えば、VBR入力ストリーム S_{in} は例えば、各々が異なるCBRビット・レートで動作する、幾つかの連続ビデオまたは時間インタバル（すなわちGOP）の連結である。各ビデオ・インタバルに対する動的ビット・レート調整が、第2.1章で*

$$\bar{C}_k = f^{-1} K F(\bar{Q}_k)$$

【0057】（バー C_{k-2} 、actual、バー Q_{k-2} 、actual）及び（バー C_{k-1} 、actual、バー Q_{k-1} 、actual）が、符号化されたばかりのGOP S_{k-2} 及び S_{k-1} の平均ビット及び平均量子化スケーラ対を表すものとする。これらのGOPの符号化のために、CBRまたはVBR-MPE G-2符号器が使用され得る。更に、バー C_k とバー Q_k ※20

$$\bar{C}_k = -\xi_k \bar{Q}_k + \eta_k$$

【0058】ここで ξ_k 及び η_k は、それぞれ次式で与えられる。

$$\xi_k = \frac{\Delta \bar{C}_k}{\Delta \bar{Q}_k}$$

【数21】

$$\eta_k = \bar{C}_{k-1, \text{actual}} + \xi_k \bar{Q}_{k-1, \text{actual}} \quad 10c)$$

【0059】更に、前記式10b)で与えられる ξ_k の分子、分母は、それぞれ次式で与えられる。 ☆【数22】

$$\Delta \bar{C}_k = \bar{C}_{k-1, \text{actual}} - \bar{C}_{k-2, \text{actual}} \quad 10d)$$

【数23】

$$\Delta \bar{Q}_k = \bar{Q}_{k-2, \text{actual}} - \bar{Q}_{k-1, \text{actual}} \quad 10e)$$

【0060】式10a)は、分析中の特定のGOPの瞬時レート・量子化動作を定義する。式10a)の線形モデル、及びそれと式9a)の知覚モデルとの関係が、図3に示される。

【0061】最適な動作ポイントを見出すために、式10a)は式9a)の知覚モデルと一緒に解かれる。例えば、理想動作ポイント3は、これらの式を（バー

$$\bar{C}_k = \frac{K(\xi_k a_1 + \eta_k a_2)}{f \xi_k + a_2 K}$$

*述べられた知覚モデルにより決定され、実際の（バー C 、バー Q ）対が式1)のレート・量子化モデルから計算される。従って、VBRシーケンス S_{in} が、 $\{S_i\}_{i=1}^{N_{in}}$ により指定される有限数のGOPに区分化される。GOP S_k に対して、平均ビット数（バー C_k ）及び平均量子化スケーラ（バー Q_k ）が、 $[\text{バー}C_k, \text{バー}Q_k] = G^{-1} [C_k, Q_k]$ となるように定義される。式9)の項が、本発明の原理に従い、次式のようにGOP S_k の平均ビットと量子化スケーラとの間の依存性を形成するように変更される。ここでビデオのフレーム・レートは f により定義される。

【数18】

$$9a)$$

※との間に、式1)で示されるような双曲型依存性が存在し、前記の対が本発明において、バー C —バー Q 関係の線形モデルを予測するために、次のように使用されるものと仮定する。

【数19】

$$10a)$$

★【数20】

$$10b)$$

☆【数22】

$$10d)$$

$$10e)$$

◆ C_1 、バー Q_1 ）及び（バー C_2 、バー Q_2 ）で解くことにより獲得され、理想動作ポイント4は、（バー C_2 、バー Q_2 ）及び（バー C_3 、バー Q_3 ）で解くことにより獲得される。

【0062】特に、次のGOPは次式の平均ビットで符号化される。

【数24】

$$11)$$

【0063】定数 K が各GOPに対して調節され、VBRストリームにより生成される総ビット数が、例えば D

VDディスクなどの記憶装置または検索装置のサイズよりも大きくならないように保証する。例えば、有効総ビット数が R_{tot} の場合、入力シーケンスのあらゆるGOP *

* Pが分析され、符号化された後、式9 a)の知覚モデルが、Kを計算するために次のように使用される。
【数25】

$$K = \frac{f R_{tot}}{G \sum_{k=0}^{N_{gop}-1} F(\bar{Q}_{k, actual})} \quad (12)$$

【0064】ここで N_{gop} は入力シーケンス内のGOPの数、GはGOPサイズである。

【0065】式12)の分母は、複数パス符号化方式において容易に計算され得るが、本発明の場合のように、単一パス・リアルタイム圧縮方式では使用可能でない。従って、本発明の原理によれば、代わりに予め符号化されたファントム・シーケンスが、次式により定義されるGOP量子化器のセットと共に使用される。

【数26】

$$\{\bar{Q}_k^*\} \quad N_{gop}^* - 1$$

$k=0,$

20

※ 【数27】

$$K_k = \frac{f R_{k, tot}}{G \sum_{k=0}^{N_{gop}-1} F(\bar{Q}_{k, actual})} \quad (13)$$

【0068】図3は、ソフトなまたはハードなGOPが符号化される場合に、知覚モデルが調整される様子を示す。

30

【0069】K、特にKの分母は、単一パスVBR符号化方式では計算され得ないので、本発明の原理に従い、Kを調節し、各GOPに割当てられるビット数を推定する技法が考案された。リアルタイム単一パスVBR方式は、その進行に伴い特定の情報を学習するので、学習プロセスと呼ばれる。学習プロセスの1実施例について、図4に関連して次に述べる。

【0070】1. Pを、次式で与えられる組により定義される予め符号化されたファントムとする。

【数28】

$$\{N_{gop}^*, \{\bar{Q}_k^*\}\} \quad N_{gop}^* - 1$$

$k=0$

【0071】そして Z_0 及び N_0 を次のように設定する(ステップ400)。

【数29】

$$Z_0 = \left(\sum_{k=0}^{N_{gop}^* - 1} F(\bar{Q}_k^*) \right)$$

【数30】 $N_0 = N_{gop}^*$

【0072】ここで Z_0 の典型値は108、 N_0 の典型値は114である。

【0073】2. 例えばCBR MPEG-2符号器を用い、入力シーケンスの最初の2つのGOP、すなわち S_0 及び S_1 を、 $f \cdot (\text{バー}C_{gop})$ のレートで符号化することにより、VBRレート制御方式を初期化する。 $f \cdot (\text{バー}C_{gop})$ の公称値は、ユーザにより与えられるVBRストリームの平均レートに対応すべきである(ステップ402)。

40

【0074】3. S_{i-1} が符号化された後、様々なパラメータが更新される(ステップ404)。例えば、符号化難度の総和を表す Z_i 、及び調整されたGOPの数 N_i が、それぞれ次のように更新される。

【数31】

$$Z_k = Z_{k-1} + \gamma F(\bar{Q}_{k-1, actual})$$

【数32】 $N_k = N_{k-1} + \gamma$

50 【0075】ここで γ は学習アルゴリズムの更新スピー

ドである（例えば $y = 2.0$ ）。更に、 $C_{j, actual}$ 及び総和 $\Sigma C_{j, actual}$ ($j = 0$ 乃至 $k-1$) が計算される。a) ここで $k-1=0$ ならば、 k を1増分し、ステップ3に移行する。b) それ以外では、式10b) 乃至式10c) により定義される予測モデルに従い、 ξ_k 及び η_k などの他のパラメータを計算する。

【0076】4. ステップ3の総和項を使用し、パジェット内の残りのビット $R_{k, tot}$ を決定する。例えば、 $\Sigma C_{j, actual}$ ($j = 0$ 乃至 $k-1$) を R_{tot} から減算することにより、 $R_{k, tot}$ を獲得する。符号化される残りのフレームまたはピクチャ P_k の数を更新する（ステップ404）。

【0077】5. 新たな傾き K_k を次式により測定する（ステップ406）。

【数33】 $K_k = f(R_{k, tot} N_k / P_k Z_k)$

【0078】6. 新たなターゲットのバー C_k を、前述 *

$$\bar{C}_k = \left(\frac{R_{k, tot} N_k}{P_k Z_k} \right) \left(\frac{\xi_k a_1 + \eta_k a_2}{\xi_k + a_2 \left(\frac{R_{k, tot} N_k}{P_k Z_k} \right)} \right) \quad (14)$$

【0083】本発明のレート制御方式は、更に図5に関連して後述される。1実施例では、非圧縮入力ストリーム（例えばビデオ・ストリーム）が圧縮エンジン500に入力される。圧縮エンジン500は、例えば図1に示され、前述された符号器を含む。圧縮エンジン500は入力ストリームの事前知識がないので、チップの古い技術、すなわち既知のCBRまたはVBRアルゴリズムにより初期化される。その後、最初の2つのGOPが、例

例えばCBRモードで符号化される。

【0084】図5では遅延502が提供され、これは情報が次に符号化されるGOPに当たるGOP_kからではなく、前に符号化されたGOP（例えばGOP_{k-1} 及びGOP_{k-2}）から獲得されることを示す。

【0085】その後、最後の2つの符号化GOP（例えばGOP_{k-1} 及びGOP_{k-2}）の各々に対する平均ビット、すなわちバー C_1 、バー C_2 、及びそれらのGOPの平均Q因子、すなわちバー Q_1 、バー Q_2 が計算される

（504）。例えば、最初の2つのGOPに対して、平均ビット及び平均Q因子が計算される様子が、式1) 乃至式7) に関連して前述された。次に、(バー C_1 、バー Q_1)、(バー C_2 、バー Q_2) の値がバッファされ、後に、符号化される次のGOP（例えばGOP_k）に対して必要とされるGOPビットを推定するために使用さ※

*の式11) または後述の式14) により計算する（ステップ408）。

【0079】7. S_k をVBRモードで符号化する（ステップ410）。

【0080】8. シーケンス内に未完のGOPが存在する場合（問い合わせ412）、ステップ3に移行し、それ以外では終了する（ステップ414）。

【0081】項 Z_k は、それ自身を各GOPのイメージ内容に適応化させ、VBRビデオの非持久性を平滑化することを目的とする。これは非常に高い（低い）複雑度のGOPに対する非現実的なビット割当てを禁ずるための安全性尺度として作用する。

【0082】以上、符号化されるGOP S_k の平均ビットを決定する方式について述べてきた。要するに、このターゲット・ビットは次の式により設定される。

【数34】

※れる。

【0086】新たなGOPビットを推定するために（506）、式10) のバー C -バー Q 予測モデル（508）、及び式9a) のバー C -バー Q 知覚モデル（510）が、新たなGOPビットの推定において使用される。図5に示されるように、学習プロシーダの更新パラメータ、消費する残りのビット、及び符号化するピクチャの数が、時変知覚モデルに入力される（511）。予測モデルと知覚モデルが一致するポイントが、GOPに対するGOPビットの推定となる。このポイントは、前述の式14) により定義される。

【0087】更に、GOPビットの推定が、後述のように特殊効果及び不安定性に対して調整され得る（512）。

【0088】新たなGOPビットが導出された後、ピクチャ・ビット割当てモデルを用いて、符号化されるGOPの各フレームまたはピクチャに対して、ビット数が割当てられる（514）。1実施例では、各ピクチャ・タイプのターゲット・ビットを調整するために、バー C_k にGOPサイズ G が乗算され、次式により各ピクチャに対する理想ビットが獲得される。

【数35】

$$R_{n, ideal}^x = \frac{X_{n-1}^x (C_k - \alpha_1 \Delta_{n-1, gop} - \alpha_2 \Delta_{n-1, pic})}{\sum_x X_{n-1}^x N^x} \quad (15)$$

【0089】CBRタイプの動作では、1 GOP 当たり同一のビット数が消費され、理想 GOP ターゲット・ビットと実際の値との間に差が存在する場合、長期 GOP 誤差が、次に符号化される GOP に持ち越される。従って、1つの GOP、GOP の集まり、またはシーケンス全体に対するビデオ・ストリームのビット・レートは同一である。長期 GOP 誤差は $\Delta_{n-1, \text{gop}}$ である。しかしながら、リアルタイム VBR 状況では、長期 GOP 誤差を持ち越す必要はない。なぜなら、各 GOP が異なるビット・レートで動作し得るからである。VBR 方式の責任は、ビデオ品質に対する総ビット・バジェットを最適化することである。前記式 15) 内のパラメータ α_1 により、長期 GOP 誤差をスケールし、この情報が VBR 符号器のレート制御アルゴリズムに供給されるときに、それを制御することが可能である。一旦 VBR の GOP ターゲット・ビットがセットされると、 α_2 が GOP 内のビット・バジェットに応えるために使用される。 α_1 及び α_2 の典型値は、それぞれ 0.05 及び 0.2 である。

【0090】例えば、CBR レートを 4.0 Mビット/秒にセットし、最初の GOP を符号化する。ここで 3.9 Mビット/秒が消費されるものと仮定する。残りの 0.1 Mビット/秒が次の GOP (番号 2) に持ち越され、従って GOP 2 は (理想的には) 4.1 Mビット/秒を獲得する。しかしながら、GOP 2 が 3.5 Mビット/秒を獲得するだけであれば、最初の 2つの GOP に対して、0.6 Mビット/秒のアンダーシュートが存在することになり、これが GOP 3 に持ち越され、区分的な一定ビット・レートを保証する。これは VBR では必要とされない。なぜなら、我々は CBR にとって弱点である領域に対して、ビデオ品質を最適化しようとしているからである。余剰ビット (0.6 Mビット/秒) を保存して、符号化が困難 (ハード) なシーンに対して、効率的に使用するべきである。これが α_1 が非常に小さな理由である。

【0091】ピクチャに対する理想ビットが決定された後、各ピクチャに対する量子化因子が、例えば式 1) を用いて導出される (516)。量子化因子は圧縮エンジン 500 に入力され、現 GOP の各フレームまたはピクチャを圧縮するために使用される。圧縮エンジンの出力 *40

$$B_n^* = B_n - R_{n, \text{actual}} \quad (16)$$

【0096】更に、次のピクチャを除去する前のバッファ充填度は、次のようになる。 ※ 【数 37】

$$B_{n+1} = \text{MIN} (B_{\text{buf}}, B_n^* + R_{\text{max}} T_n) \quad (17)$$

【0097】ここで B_{buf} はビデオ・バッファ検証器の合計サイズで、 T_n はピクチャ n の表示期間である (T_n は f の逆数)。ピクチャ n が除去される前は、バッファ内の全ての情報が使用可能であり、従って上限 U_n がバ

$$U_n = \text{MIN} (B_{\text{buf}}, B_{n+1} + R_{\text{max}} T_{n+1} - R_{n+1, \text{actual}}) \quad (18a)$$

【数 39】

* は、本発明の原理に従い、単一パス・リアルタイム可変ビット・レート・モードで圧縮されたストリームである。

【0092】フレームの符号化に続き、各ピクチャに割当てられた実際のビット数が、例えば圧縮エンジン 500 内のカウンタから獲得される。これらの値が式 6) に挿入され、実際のビットが理想ビットからどの程度かけ離れているかを決定する。

【0093】図 5 では、GOP ビットの数推定するために使用される様々なプロセスが、圧縮エンジン 500 とは別々に示されている。すなわち、圧縮エンジン 500 とは独立の処理論理または回路が、プロセスを実現するために使用され得る。しかしながら、別の実施例では、図 5 の 1 つ以上のプロセスが 1 チップ上に (すなわち圧縮エンジン 500 内に) 含まれる。

【0094】前述の単一パス VBR レート・量子化フレームワークは、伸張エンジン内の復号器バッファが無限のサイズであり、十分な数のビットが常に使用可能であるといった仮定にもとづく。しかしながら、実際の状況では、バッファ・サイズは MPEG-2 規格により制限され、定義される。従って、本発明の符号化方式は、復号器バッファが遭遇し得るオーバフロー条件またはアンダーフロー条件を排除する責任がある。これは仮想的な復号器バッファ、すなわちビデオ・バッファ検証器 (V BV) を調査し、ピクチャ・タイプに割当てられるビットの量の上限及び下限を計算することにより達成される。幸運にも、VBR 状況では、復号器バッファにおいてオーバフローは発生し得ない。なぜなら、V BV バッファ占有がその最大レベルに達したとき、復号器バッファを充填するタスクが、即時停止されるからである。従って、ピクチャ限界に対して、0 の下限が課せられる。更に、復号器バッファが、ユーザにより設定された R_{max} (Mビット/秒) の最大レートで充填される。

【0095】しかしながら、VBR 復号器の V BV において、アンダーフロー条件は発生し得、所定値以下の上限がこれを避けるために割当てられるべきである。従って、 B_n 及び B_n^* を、それぞれピクチャ n が除去される前の、及び後の復号器バッファ充填度とする。 B_n^* は次のように計算される。

※ 【数 36】

※ 【数 37】

※ 【数 38】

※ 【数 39】

※ 【数 40】

※ 【数 41】

※ 【数 42】

※ 【数 43】

※ 【数 44】

※ 【数 45】

※ 【数 46】

※ 【数 47】

※ 【数 48】

※ 【数 49】

※ 【数 50】

$$L_n = 0$$

18b)

【0098】従って、理想ピクチャ・ビットは、次のピクチャ限界を用いてクリップされる。 * 【数40】

$$R_{n, ideal} = \begin{cases} U_n & \text{if } (R_{n, ideal} > U_n) \\ L_n & \text{elseif } (R_{n, ideal} < L_n) \\ R_{n, ideal} & \text{else} \end{cases} \quad (19)$$

【0099】これまで述べてきたのは、 $\{R_m, Q_m\}_{m=0}^{n-1}$ により指定される単一パスVBR MPEG-2符号器である。符号器は、時変知覚モデルと、(パーC-パーQ)モデルのバンクとを連带的に解くことにより形成されるポイント集合に従って動作する。最初に、このポイント集合内におけるGOPの(パーC₁、パーQ₁)対の相対理想位置が、GOPの"ソフトネス"及び"ハードネス"により決定され、次にビット・バジェットの残りのビット数により調整される。ピクチャの(R₁、Q₁)対のローカル位置は、ピクチャ・タイプの(R-Q)モデルにより表される。図3は、(パーC₁、パーQ₁)対のポイント集合が、可変ビット・レート制御方式1を使用するVBR符号器により形成される様子を示す。この図において、パーC=パーC_{opt}により定義される一定の直線は、全ての(パーC₁、パーQ₁)対がCBRモードで符号化されると仮定した場合の、それらの理想位置を示す。更に図3は、GOPのハードネスが増加すると、平均量子化スケアラが増加することを反映している。理想及び実際の動作ポイントが、それぞれ白丸及び黒丸により示される。(パーC₁、パーQ₁)及び(パーC₂、パーQ₂)を、ビデオ・シーケンスの最初に、VBR符号器の初期化により獲得される"ソフト"GOPの最初の2つの対とする。"3"で示される次の理想動作ポイントは、知覚モデル(実線)と因果予測モデルとの交点である。ポイント"3"の位置は、シーケンス平均ビットすなわちパーC_{opt}よりも、小さな平均ビットが割当てられることを示す。しかしながら、符号器の出力すなわちポイント(パーC₁、パーQ₁)は異なり、その結果、ビット・バジェット制約に適合する新たな知覚モデル(点線で示される)が導出される。次に、(パーC₂、パーQ₂)及び(パーC₃、パーQ₃)の対が、新たな予測モデル(点線で示される)を獲得し、新たな動作ポイントすなわち"4"を計算するために使用される。ビデオ・シーケンスが多数の連続する"ソフト"GOPを含む場合、知覚モデルは最終的には、CBR出力対(パーC₁、パーQ₁)及び(パーC₂、パーQ₂)の近くでパーC=パーC_{opt}の直線と交差する直線に収束する。

【0100】ビデオ資料が"ハード"GOPで開始する場合、実際のポイント(パーC₃、パーQ₃)の位置は、パーC=パーC_{opt}の直線よりも上にある。このことはCBR符号器が最も弱い領域、すなわち"ハード"GOPにおいて使用可能なビット・バジェットの最適化を保証

る。この状況では、新たな知覚モデル(破線により示される)が、割当てられるGOPビットを次第に少なくしてビット・バジェット制約に適合するように形成される。このモデルは、新たな因果予測モデル(破線で示される)と共に、"4"で示される次の動作ポイントを決定する。入来シーケンスが"ハード"GOPだけから成る場合、知覚モデルは最終的には、(パーC₁、パーQ₁)ポイント及び(パーC₂、パーQ₂)ポイントの近くの位置で、パーC=パーC_{opt}の直線と交差する直線に従う。通常のビデオ・プログラムでは、長い期間同一のGOP(または前に計算されたポイントの集合)で動作することはありそうもないが、短時間の後に、あるGOPから隣接のGOPにジャンプすることはあり得る。従って、次の動作を推論することができる。知覚モデルが、GOPの"ハードネス"または"ソフトネス"を単調に減じ得る状況に落ち着く前に、レート制御方式は新たなGOPに移動する。前に符号化されたピクチャ・タイプの実際の量子化スケアラ値が、"よりハードな"GOPまたは"よりソフトな"GOPへの移動を決定する。図3に示されるように、将来の"よりハードな"("よりソフトな")GOPの場合は、前に符号化されたGOPの右(左)に移動する。従って、知覚モデルと交差するところで、前のGOPにより使用された実際のビットよりも高い(低い)GOPビットを生成する。この機構は、類似の"符号化難度"を有する連続するGOPの集合に対して、"ハードネス"(または"ソフトネス")の減少または増大の効果が、対応するビデオ・インタバルの符号化内容全体に渡って分散される一方で、最も複雑度の高い(低い)GOPが依然、その類似のGOPに対して、最大(最小)のビット割当てを獲得することを保証する。この技法は知覚モデルの傾きKの調節にもとづくが、別のVBRレート制御方式は、ユーザによりセットされた総ビット数に合うように、知覚モデルを変換することにより公式化され得る。この可変ビット・レート制御方式について、次に述べる。

【0101】2.3 VBRレート制御方式2: この章では、符号化される各GOPのターゲット・ビットを計算する第2の方式について述べる。本発明の原理によれば、(パーC-パーQ)知覚モデルの傾きKを調節する代わりに、パラメータa₁について解くことにより、知覚モデルの位置が変換される。傾きKは、f・(パーC_{opt})のCBRレートに固定され、パラメータa₁が次のように導出される。

【数41】

$$a_1 = \frac{R_{\text{TOT}} - C_{\text{GOP}} a_2 \sum_{k=0}^{N_{\text{GOP}}-1} \frac{1}{Q_k}}{C_{\text{GOP}} N_{\text{GOP}}} \quad (20)$$

【0102】前記式の総和項は、リアルタイム単一パスVBR符号器では未定義である。従って、予め符号化されたファントム・シーケンス（方式1と同様）が、本発明のVBRレート制御方式2を初期化するために使用される。デジタル記憶媒体がオーバーランまたはアンダー*10

*ランに遭遇しないことを保証するために、残りのビットが周期的にモニタされ、瞬時総ビット・バジェットすなわち $R_{i, \text{tot}}$ として使用される。知覚モデルの変換は、次のように変化する。

【数42】

$$a_1^k = \frac{P_k^{-1} R_{k, \text{tot}}}{C_{\text{GOP}}} - a_2 \frac{Y_k}{N_k} = r_a^k - a_2 E_k \quad (21)$$

【0103】ここで項 r_a^k は、まだ符号化されてないGOPの長期レートと、ビデオ・ストリームのCBRレートとの比率を表し、 Y_k 及び E_k は後述される。ハードGOPでは（バー $C_i >$ バー C_{GOP} ）であり、従って、比率 r_a^k は総ビット・バジェットを満足するために、時間と共に小さくならなければならない。更に、（バー C_i 、*20

*バー Q_k ）対の位置をポイント集合に沿って下流に移動し、圧縮ストリームのレートを下げる強制が、 E_k により提供される。幾つかのソフトGOPが時間と共に符号化される場合には、反対の状況が発生する。各GOPの平均ビットが次のように設定される。

【数43】

$$\bar{C}_k = \frac{\xi_k a_1^k + \eta_k a_2}{\xi_k \bar{C}_{\text{GOP}}^{-1} + a_2} \quad (22)$$

【0104】本発明の原理によれば、学習プロシージャがVBRレート制御方式2に対しても提供される。この学習プロシージャの間、 Y が初期化され、更新される。残りの符号化パラメータは、VBRレート制御方式1の公式化により定義される。

【0105】学習プロシージャ2の1実施例が、図6に関連して次に述べられる。

【0106】1. P を、次式で与えられる組により定義される予め符号化されたファントムとする。

【数44】

$$\{N_{\text{GOP}}^*, \{\bar{Q}_k^*\} \}_{k=0}^{N_{\text{GOP}}-1}\}$$

【0107】そして、 Y_0 及び N_0 を次のように設定する（ステップ600）。

【数45】

$$Y_0 = \left(\sum_{k=0}^{N_{\text{GOP}}-1} \frac{N_{\text{GOP}}^*}{\bar{Q}_k^*} \right)$$

【数46】 $N_0 = N_{\text{GOP}}^*$

【0108】ここで Y_0 の典型値は1869である。

【0109】2. 例えばCBR-MPEG-2符号器を用い、ビデオ・シーケンスの最初の2つのGOPすなわち S_0 及び S_1 を、 $f \cdot$ （バー C_{GOP} ）のレートで符号化することにより、VBRレート制御方式を初期化する。

$f \cdot$ （バー C_{GOP} ）の公称値は、VBRストリームの平均レートに対応すべきである（ステップ602）。

【0110】3. S_{k+1} が符号化された後、様々なパラメータが更新される（ステップ604）。例えば、フレーム・グループ当たりの平均量子化因子の総和を表す Y_k 、及び N_k が、それぞれ次のように更新される。

【数47】

$$Y_k = Y_{k-1} + \gamma \bar{Q}_{k-1, \text{actual}}$$

【数48】 $N_k = N_{k-1} + \gamma$

【0111】ここで γ は学習アルゴリズムの更新スピードである（例えば $\gamma = 2.0$ ）。更に、 $C_{k-1, \text{actual}}$ 、次に総和 $\Sigma C_{j, \text{actual}}$ （ $j = 0$ 乃至 $k-1$ ）が計算される。

a) ここで $k-1 = 0$ ならば、 k を1増分し、ステップ3に移行する。b) それ以外では、式10b)乃至式10c)により定義される予測モデルに従い、パラメータ ξ_k 及び η_k を計算する。

【0112】4. ステップ3の総和項を使用し、バジェット内の残りのビット $R_{i, \text{tot}}$ を決定する。符号化される残りのピクチャ P_k の数を更新する（ステップ604）。

【0113】5. r_a^k 、 E_k 及び新たな変換因子 a_1^k を計算する（ステップ606）。

【0114】6. 新たなターゲット、バー C_i を、例え

ば式22)により計算する(ステップ608)。

【0115】7. S_kをVBRモードで符号化する(ステップ610)。

【0116】8. シーケンス内に未完のGOPが存在する場合(問い合わせ612)、ステップ3に移行する。それ以外では、終了する(ステップ614)。

【0117】VBRバッファ占有及びターゲット調整に対する制約が、前章で述べられたように処理される。図7は、瞬時因果予測モデルが各GOP内で形成され、時変知覚モデルと共に、可変ビット・レート制御方式2を使用するVBR符号器の、新たな(バーC_k、バーQ_k)ポイントを推定するために使用される様子を示す。幾つかのソフト(またはハード)GOPが時間と共に符号化される場合、知覚モデルはビット・バジェット制約に適合するために、上方(または下方)にシフトされる。

【0118】2. 4 シーン遷移: 符号化プロシーダの間、VBRレート制御方式は不安定になり得る。この原因は通常、特殊効果または自然に発生する現象、例えば、シーン・カット、低速移動フェードまたは高速移動フェード(例えば旧シーンから暗転を経て新シーンへのフェード)、及び突然の輝度変化(カメラからの高輝度*

*な光、映画の発砲などの結果)などが、圧縮ビデオ・パラメータの非持久的性質と関連し合うことによる。前述の時間遷移及び空間遷移では、符号器のパラメータ調整がほとんど不可能となる。

【0119】本発明の原理によれば、たとえ特殊効果が圧縮されるときでも、高品質のビデオVBRストリームが提供される。これはVBR符号器の(特殊効果の結果としての)アルゴリズム的不安定性を検出し、GOPターゲットを前に述べたプロシーダと異なる軌道に沿って設定することにより達成される。

【0120】図8は、ソフトGOPからハードGOPへの遷移が、符号化される将来のGOPに対して、不正な推定を行う場合を示す。この不正な推定は図中、“不良動作ポイント”で示される。より良好な推定が、慎重に選択されたパラメータのセットに、3ポイント・メディアン(MEDIAN)・フィルタを適用することにより実行され得る。

【0121】例えば、VBRレート制御方式1において、新たなターゲットが次式により導出される。

【数49】

23)

$$C_{k,new} = G \cdot \begin{cases} \text{MEDIAN}(\mu_1 f^{-1} K_k F(\bar{Q}_{k-1,actual}), \mu_2 \bar{C}_{gop}, \mu_3 f^{-1} K_k) & \text{if } (\xi_k < 0) \\ \bar{C}_k & \text{else} \end{cases}$$

【0122】別の例として、VBRレート制御方式2において、新たなターゲットが次式により導出される。 ※ 【数50】

24)

$$C_{k,new} = G \cdot \begin{cases} \text{MEDIAN}(\mu_1 f^{-1} K_k F(\bar{Q}_{k-1,actual}), \mu_2 \bar{C}_{gop}, \mu_3 f^{-1} K_k) & \text{if } (\xi_k < 0) \\ \bar{C}_k & \text{else} \end{cases}$$

【0123】ここで $H(k) = a_1 (N_{gop}^* + \gamma k) N_k^{-1}$ である。

【0124】メディアン・フィルタは、全てのタイプのGOP、例えばソフトGOPまたはハードGOP内で発生し得る、(シーン遷移の結果としての)不安定条件を補償するために使用される。定数セット $\{\mu_1, \mu_2, \mu_3\}$ が、シーン遷移の間に生成されるビット数に対して限界を強制する。例えば、 $\mu_1 = 1.0$ 、 $\mu_2 = 0.9$ 、 $\mu_3 = 0.9$ である。

【0125】式23)及び式24)のGOPターゲット割当ては、シーン変化が検出されるときにも使用され得る。シーン・カットの検出方法については、多くの技術が文献などで述べられているのでここでは述べない。各GOPの符号化はIピクチャにより開始するので、このピクチャの位置がシーケンス内でシーン・カットが発生する位置に相当する場合、新たなターゲットが式23)

または式24)により設定される。更に、符号器がPピクチャに対してシーン変化を検出する場合、このタイプのピクチャがIタイプに変更されて、新たなGOPが開始され、再度、新たなターゲットが式23)または式24)に従い設定される。

【0126】2. 5 (バーCーバーQ)不安定性: 式1)の(R-Q)関係は、多様なクラスのビデオ・シーケンスに対して、かなり良好に作用することが知られており、多くのリアルタイムMPEG圧縮システム及び非MPEG圧縮システムにおいて使用される基本概念である。しかしながら、本発明の原理によれば、この関係は、先行GOPの実際の平均GOPビットと平均量子化因子にもとづき、符号化される現GOPの平均ビットを推定するための、因果予測モデルを生成するために使用される。しかしながら、この予測モデルは、たとえ特殊効果、シーン遷移またはシーン・カットが存在しない場

合でも、特定のビデオ・シーケンスに対して信頼できないものになる。この原因は、ビデオ・ソースの非定常的な性質、例えば、大きな輝度またはクロミナンスのイメージ詳細の背景中をゆっくり移動する幾つかの高度に詳細なオブジェクト、或いはシーン内の不所望な雑音の存在などに関係する。

【0127】図9は、符号器により導出される2つのポイントの一方が、合成（バーCーバーQ）曲線の外側に*

* 存在する場合を示す。このポイントは、前述の基準の結果であり、（バーCーバーQ）関係に従わない。更に、これは“不良動作ポイント”の計算に寄与する。本発明のVBR符号器は、式23）または式24）で定義される方法に対して、追加の寄与を行うことにより、この不安定性を抑制する。特に、GOPビットが下記の条件に従い設定される。

【数51】

25)

$$C_{k, new} = \begin{cases} MAX(C_{k, new}, G\Phi_{max} \bar{C}_{k, fill}) & \text{if } (\xi_k < 0) \wedge (\theta_L < |\Delta \bar{Q}| < \theta_U) \wedge (\bar{C}_{k, fill} > \bar{C}_{gop}) \\ MIN(C_{k, new}, G\Phi_{min} \bar{C}_{k, fill}) & \text{elseif } (\xi_k < 0) \wedge (\theta_L < |\Delta \bar{Q}| < \theta_U) \wedge (\bar{C}_{k, fill} < \bar{C}_{gop}) \\ G\bar{C}_{gop} & \text{elseif } (\xi_k < 0) \wedge (\theta_L < |\Delta \bar{Q}| < \theta_U) \wedge (\bar{C}_{k, fill} = \bar{C}_{gop}) \end{cases}$$

【0128】ここでバーC、fillは次式で与えられる。 ※ ※ 【数52】

$$\bar{C}_{k, fill} = \frac{\omega_1 \bar{C}_{k-2, actual} + \omega_2 \bar{C}_{k-1, actual}}{\omega_1 + \omega_2} \quad 26)$$

【0129】ここで ω_1 及び ω_2 はフィルタの重みを表す。1例では、 $\omega_1 = 1, 0$ 、 $\omega_2 = 1, 0$ であり、別の例では、 $\omega_1 = 1, 0$ 、 $\omega_2 = 3, 0$ に設定される。

【0130】前記式25）では、 Φ は割当てられるビット数を調節するために使用される定数である。例えば、 $MAX \Phi = 1, 1$ 、 $MIN \Phi = 0, 9$ である。しきい値 θ_L （例えば0, 1）及び θ_U （例えば1, 0）は、将来のGOPが既に処理済みのGOPと同一のビデオ・セグメントに属するとの仮定の下で選択され、バーC、fillは、前に符号化されたGOPのビットに線形フィルタを適用することにより決定される、平均GOPビットである。この線形フィルタの出力は、新たな動作ポイントである。MAX関数及びMIN関数は、（バーCーバーQ）不安定性を生じる符号化ビット・パラメータを除去するために実行される。

【0131】これまで、容易に実現でき、高品質のビデオを生成するリアルタイム単一パス可変ビット・レート符号器を提供する技術について述べてきた。GOPに割当てられるビット数が推定され、次にその推定が様々な技術を用いて改善される。

【0132】本発明は、例えばコンピュータ使用可能媒体を有する製造物（例えば1つ以上のコンピュータ・プログラム製品）内に含まれ得る。媒体は例えば、本発明の能力を提供し、容易にするためのコンピュータ読出し可能プログラム・コード手段を実現する。そのような製造物はコンピュータ・システムの一部として含まれるか、別々に販売される。

【0133】更に、本発明の能力を実行するために、マシンにより実行可能な少なくとも1つの命令プログラムを実現する、マシンにより読出し可能な少なくとも1つのプログラム記憶装置が提供される。

【0134】ここで示されたフロー図は模範例に過ぎない

い。これらの図またはステップ（または動作）には、本発明の趣旨から逸れることなく多くの変化が存在し得る。例えば、ステップが異なる順序で実行されたり、ステップが追加されたり、消去または変更され得る。これらの全ての変化は、本発明の一部と見なされる。

【0135】まとめとして、本発明の構成に関して以下の事項を開示する。

【0136】（1）フレーム・グループを処理する方法であって、入力ストリームの少なくとも1つのフレーム・グループを符号化するステップと、前記入力ストリーム全体を符号化する前に、前記符号化するステップから獲得される1つ以上のパラメータ値を用い、可変ビット・レート・モードで符号化される、前記入力ストリームの別のフレーム・グループに割当てられるビット数を決定するステップとを含む、方法。

（2）前記別のフレーム・グループを、単一パス可変ビット・レート・モードで、1つの符号化パスだけにより符号化するステップを含む、前記（1）記載の方法。

（3）前記入力ストリームの他のフレーム・グループに対して、前記決定するステップ及び前記別のフレーム・グループを符号化するステップを繰り返すステップを含み、前記他のフレーム・グループの各々に対して決定するステップが、1つ以上の前に符号化されたフレーム・グループから獲得される1つ以上のパラメータ値を使用する、前記（2）記載の方法。

（4）前記別のフレーム・グループを符号化するステップが、前記別のフレーム・グループの1つ以上のフレームに割当てられる理想ビット数を決定するステップと、復号器バッファにおいて、オーバーフローまたはアンダーフローが生じないように、前記理想ビット数を調整するステップとを含む、前記（2）記載の方法。

（5）前記調整するステップが、前記1つ以上のフレ

ムに割当てられる前記理想ビット数を制限するために使用される、1つ以上のピクチャ限界を決定するステップを含む、前記(4)記載の方法。

(6) 前記決定するステップが、前記別のフレーム・グループに対して、所定の知覚関数の傾きを調節するステップと、調節された傾きをビット・レート方程式内で用い、前記ビット数を決定するステップとを含む、前記(1)記載の方法。

(7) 前記所定の知覚関数が、

【数53】

$$\bar{C}_k = \left(\frac{R_{k, \text{tot}} N_k}{P_k Z_k} \right)^{*10} \left(\frac{\xi_k a_1 + \eta_k a_2}{\xi_k + a_2 \left(\frac{R_{k, \text{tot}} N_k}{P_k Z_k} \right)} \right)$$

で与えられ、ここで $R_{k, \text{tot}}$ はビット・パケット内の残りのビット数、 N_k は入力ストリーム内のフレーム・グループの調整された数、 P_k は符号化される前記別のフレーム・グループの残りのフレームの数、 Z_k は符号化難度の総和をそれぞれ表し、 $\xi_k = \Delta (\text{バー} C_k) / \Delta (\text{バー} Q_k)$ であり、 $\text{バー} C_k$ は平均ビット、 $\text{バー} Q_k$ は平均量子化因子を表し、 $\eta_k = \text{バー} C_{k-1, \text{actual}} + \xi_k$ 、 $(\text{バー} Q_{k-1, \text{actual}})$ であり、 a_1 及び a_2 が定数である、前記(6)記載の方法。

(9) 前記決定するステップが、前記別のフレーム・グループに対して、所定の知覚関数の変換因子を決定するステップと、前記変換因子をビット・レート方程式内で使用し、前記ビット数を決定するステップとを含む、前記(1)記載の方法。

(10) 前記変換因子を決定するステップが、

$$a_1^k = \frac{P_k^{-1} R_{k, \text{tot}}}{C_{\text{top}}} - a_2 \frac{Y_k}{N_k}$$

を使用し、ここで P_k は符号化される前記別のフレーム・

で与えられ、ここで $R_{k, \text{tot}}$ はビット・パケット内の残りのビット数、 N_k は入力ストリーム内のフレーム・グループの調整された数、 P_k は符号化される前記別のフレーム・グループの残りのフレームの数、 Z_k は符号化難度の総和をそれぞれ表し、 $\xi_k = \Delta (\text{バー} C_k) / \Delta (\text{バー} Q_k)$ であり、 $\text{バー} C_k$ は平均ビット、 $\text{バー} Q_k$ は平均量子化因子を表し、 $\eta_k = \text{バー} C_{k-1, \text{actual}} + \xi_k$ 、 $(\text{バー} Q_{k-1, \text{actual}})$ であり、 a_1 及び a_2 が定数である、前記(12)記載の方法。

*

$$\bar{C}_k = f^{-1} K F (\bar{Q}_k)$$

で与えられ、ここで $F (\text{バー} Q_k) = a_1 + a_2 (\text{バー} Q_k / \beta)$ であり、 a_1 、 a_2 、 β は定数、 $\text{バー} Q_k$ は量子化因子、及び K は傾きの値をそれぞれ表す、前記(6)記載の方法。

(8) 前記ビット・レート方程式が、

【数54】

$$\bar{C}_k = \left(\frac{R_{k, \text{tot}} N_k}{P_k Z_k} \right)^{*10} \left(\frac{\xi_k a_1 + \eta_k a_2}{\xi_k + a_2 \left(\frac{R_{k, \text{tot}} N_k}{P_k Z_k} \right)} \right)$$

※・グループの残りのフレームの数、 $R_{k, \text{tot}}$ はビット・パケット内の残りのビット数、 $\text{バー} C_{k-1}$ は前記入力ストリームのレート、 Y_k は1フレーム・グループ当たりの平均量子化因子の総和、 N_k は入力ストリーム内のフレーム・グループの調整された数をそれぞれ表し、 a_1 は定数である、前記(9)記載の方法。

(11) 前記符号化するステップが、前記入力ストリームの第1のフレーム・グループ及び第2のフレーム・グループを符号化するステップを含み、前記1つ以上のパラメータ値が、前記第1のフレーム・グループに対する第1のビット・レート及び第1の量子化因子と、前記第2のフレーム・グループに対する第2のビット・レート及び第2の量子化因子とを含む、前記(1)記載の方法。

(12) 前記決定するステップが、前記第1のビット・レート、前記第2のビット・レート、前記第1の量子化因子、及び前記第2の量子化因子を用いる所定の関数により、割当てられる前記ビット数を決定するステップを含む、前記(11)記載の方法。

(13) 前記所定の関数が、

【数56】

$$\bar{C}_k = \left(\frac{R_{k, \text{tot}} N_k}{P_k Z_k} \right)^{*10} \left(\frac{\xi_k a_1 + \eta_k a_2}{\xi_k + a_2 \left(\frac{R_{k, \text{tot}} N_k}{P_k Z_k} \right)} \right)$$

(14) フレーム・グループを処理する方法であって、1つ以上の前に符号化されたフレーム・グループから、複数のパラメータ値を決定するステップと、単一パス可変ビット・レート・モードで符号化されるフレーム・グループに対して、前記複数のパラメータ値の1つ以上を用い、所定の関数の傾きを計算するステップと、計算された傾き及び前記1つ以上のパラメータ値を用い、符号化される前記フレーム・グループのビット・レートを獲得するステップとを含む、方法。

(15) 前記可変ビット・レート・モードで符号化される前記フレーム・グループを、当該グループに対して獲得された前記ビット・レートで符号化するステップを含む、前記(14)記載の方法。

(16) 前記入力ストリームの他のフレーム・グループに対して、前記決定するステップ、前記計算するステップ、前記獲得するステップ、及び前記符号化するステップを繰り返すステップを含む、前記(15)記載の方法。

(17) 前記傾きを計算するステップが、次の関数

$$\overline{C}_k = \left(\frac{R_{k, \text{tot}} N_k}{P_k Z_k} \right)$$

で与えられ、ここで $R_{k, \text{tot}}$ はビット・パケット内の残りのビット数、 N_k は入力ストリーム内のフレーム・グループの調整された数、 P_k は符号化される前記別のフレーム・グループの残りのフレームの数、 Z_k は符号化難度の総和をそれぞれ表し、 $\xi_k = \Delta(\text{バー}C_k) / \Delta(\text{バー}Q_k)$ であり、 $\text{バー}C_k$ は平均ビット、 $\text{バー}Q_k$ は平均量子化因子を表し、 $\eta_k = \text{バー}C_{k-1, \text{actual}} + \xi_k$ 、 $\text{バー}Q_{k-1, \text{actual}}$ であり、 a_1 及び a_2 が定数である、前記(18)記載の方法。

(20) 前記所定の関数が、前記別のフレーム・グループの瞬時レート・量子化動作、及び前記別のフレーム・グループの複雑度を示す知覚モデルを表わす、前記(18)記載の方法。

(21) 前記レート・量子化動作が、

$$\overline{C}_k = -\xi_k \overline{Q}_k + \eta_k$$

により定義され、ここで $\text{バー}C_k$ は平均ビット、 $\text{バー}Q_k$ は平均量子化スケラ、 $\xi_k = \Delta(\text{バー}C_k) / \Delta(\text{バー}Q_k)$ であり、 $\Delta(\text{バー}C_k) = \text{バー}C_{k-1, \text{actual}} - \text{バー}C_{k-2, \text{actual}}$ 、 $\Delta(\text{バー}Q_k) = \text{バー}Q_{k-1, \text{actual}} - \text{バー}Q_{k-2, \text{actual}}$ であり、 $\eta_k = \text{バー}C_{k-1, \text{actual}} + \xi_k$ 、 $\text{バー}Q_{k-1, \text{actual}}$ である、前記(20)記載の方法。

(22) 前記知覚モデルが、

$$\overline{C}_k = f^{-1} K F(\overline{Q}_k)$$

により定義され、ここで $F(\text{バー}Q_k) = a_1 + a_2(\text{バー}Q_k)^\beta$ であり、 a_1 、 a_2 、 β は定数、 $\text{バー}Q_k$ は量子化因子、及び K は傾きの値をそれぞれ表す、前記(20)記載の方法。

(23) フレーム・グループを処理する方法であって、入力ストリームの1つのフレーム・グループに割当てら

*を使用し、ここで f は前記入力ストリームのフレーム・レート、 $R_{k, \text{tot}}$ はビット・パケット内の残りの数、 N_k は入力ストリーム内のフレーム・グループの調整された数、 P_k は符号化される前記別のフレーム・グループの残りのフレームの数、 Z_k は符号化難度の総和をそれぞれ表わす、前記(14)記載の方法。

(18) 前記獲得するステップが、所定の関数を用い、前記ビット・レートを獲得するステップを含む、前記(14)記載の方法。

(19) 前記所定の関数が、

$$\left(\frac{\xi_k a_1 + \eta_k a_2}{\xi_k + a_2 \left(\frac{R_{k, \text{tot}} N_k}{P_k Z_k} \right)} \right)$$

れるビット数を、前記フレーム・グループの瞬時レート・量子化動作、及び前記フレーム・グループの複雑度を示す知覚モデルを用い、決定するステップと、前記フレーム・グループを可変ビット・レート・モードにより、前記決定するステップにより定義されたレートで符号化するステップとを含む、方法。

(24) 単一パス可変ビット・レート符号器であって、入力ストリームの少なくとも1つのフレーム・グループを符号化するように適応化された符号化エンジンと、前記入力ストリーム全体を符号化する前に、前記少なくとも1つのフレーム・グループの符号化から獲得された1つ以上のパラメータ値を用い、可変ビット・レート・モードで符号化される、前記入力ストリームの別のフレーム・グループに割当てられるビット数を決定する手段とを含む、符号器。

(25) 前記別のフレーム・グループを符号化するために前記符号化エンジンが、前記別のフレーム・グループの1つ以上のフレームに割当てられる理想ビット数を決定する手段と、復号器バッファにおいて、オーバーフローまたはアンダーフローが生じないように、前記理想ビット数を調整する手段とを含む、前記(24)記載の符号器。

(26) 前記調整する手段が、前記1つ以上のフレームに割当てられる前記理想ビット数を制限するために使用される、1つ以上のピクチャ限界を決定する手段を含む、前記(25)記載の符号器。

(27) 前記決定する手段が、前記別のフレーム・グループに対して、所定の知覚関数の傾きを調節する手段と、調節された傾きをビット・レート方程式内で用い、前記ビット数を決定する手段とを含む、前記(24)記載の符号器。

(28) 前記決定する手段が、前記別のフレーム・グループに対して、所定の知覚関数の変換因子を決定する手段と、前記変換因子をビット・レート方程式内で用い、

20

30

40

50

前記ビット数を決定する手段とを含む、前記(24)記載の符号器。

(29) 前記少なくとも1つのフレーム・グループが、前記入力ストリームの第1のフレーム・グループ及び第2のフレーム・グループを含み、前記1つ以上のパラメータ値が、前記第1のフレーム・グループに対する、第1のビット・レート及び第1の量子化因子と、前記第2のフレーム・グループに対する、第2のビット・レート及び第2の量子化因子とを含む、前記(24)記載の符号器。

(30) 前記決定する手段が、前記第1のビット・レート、前記第2のビット・レート、前記第1の量子化因子、及び前記第2の量子化因子を用いる所定の関数により、割当てられる前記ビット数を決定する手段を含む、前記(29)記載の符号器。

(31) 単一パス可変ビット・レート符号器であって、1つ以上の前に符号化されたフレーム・グループから、複数のパラメータ値を決定する手段と、単一パス可変ビット・レート・モードで符号化されるフレーム・グループに対して、前記複数のパラメータ値の1つ以上を用い、所定の関数の傾きを計算する手段と、計算された傾き及び前記1つ以上のパラメータ値を用い、符号化される前記フレーム・グループのビット・レートを獲得する手段とを含む、符号器。

(32) 前記可変ビット・レート・モードで符号化される前記フレーム・グループを、当該グループに対して獲得された前記ビット・レートで符号化する手段を含む、前記(31)記載の符号器。

(33) 前記獲得する手段が、所定の関数を用い、前記ビット・レートを獲得する手段を含む、前記(31)記載の符号器。

(34) 単一パス可変ビット・レート符号器であって、入力ストリームの1つのフレーム・グループに割当てられるビット数を、前記フレーム・グループの瞬時レート・量子化動作、及び前記フレーム・グループの複雑度を示す知覚モデルを用い、決定する手段と、前記フレーム・グループを可変ビット・レート・モードにより、前記決定する手段により定義されたレートで符号化するように適応化された符号化エンジンとを含む、符号器。

(35) フレーム・グループの処理をコンピュータに指示するコンピュータ読出し可能プログラム・コード手段を有する、少なくとも1つのコンピュータ使用可能媒体を含む装置であって、前記コンピュータ読出し可能プログラム・コード手段が、入力ストリームの少なくとも1つのフレーム・グループを符号化するように、前記コンピュータに指示する手段と、前記入力ストリーム全体を符号化する前に、前記符号化から獲得された1つ以上のパラメータ値を用い、可変ビット・レート・モードで符号化される、前記入力ストリームの別のフレーム・グループに割当てられるビット数を決定するように、前記コ

ンピュータに指示する手段とを含む、装置。

(36) 前記別のフレーム・グループを単一パス可変ビット・レート・モードで1つの符号化パスだけにより符号化するように、前記コンピュータに指示する手段を含む、前記(35)記載の装置。

(37) 前記ビット数を決定するように前記コンピュータに指示する手段が、前記別のフレーム・グループに対して、所定の知覚関数の傾きを調節するように、前記コンピュータに指示する手段と、調節された傾きをビット・レート方程式内で用い、前記ビット数を決定するように、前記コンピュータに指示する手段とを含む、前記(35)記載の装置。

(38) 前記ビット数を決定するように前記コンピュータに指示する手段が、前記別のフレーム・グループに対して、所定の知覚関数の変換因子を決定するように、前記コンピュータに指示する手段と、前記変換因子をビット・レート方程式内で用い、前記ビット数を決定するように、前記コンピュータに指示する手段とを含む、前記(35)記載の装置。

(39) 前記ビット数を決定するように前記コンピュータに指示する手段が、所定の関数を用い、前記ビット数を決定するように、前記コンピュータに指示する手段を含む、前記(35)記載の装置。

(40) フレーム・グループの処理をコンピュータに指示するコンピュータ読出し可能プログラム・コード手段を有する、少なくとも1つのコンピュータ使用可能媒体を含む装置であって、前記コンピュータ読出し可能プログラム・コード手段が、入力ストリームの1つのフレーム・グループに割当てられるビット数を、前記フレーム・グループの瞬時レート・量子化動作、及び前記フレーム・グループの複雑度を示す知覚モデルを用い、前記コンピュータに決定するように指示する手段と、前記フレーム・グループを可変ビット・レート・モードにより、前記手段により定義されたレートで符号化するように、前記コンピュータに指示する手段とを含む、装置。

【図面の簡単な説明】

【図1】本発明のリアルタイム単一パス可変ビット・レート制御方法を組み込み、使用するMPEG-2符号器のブロック図の1例を示す図である。

【図2】Pピクチャ・タイプ及びBピクチャ・タイプの動き推定のブロック図の1例を示す図である。

【図3】バーカーQモデルの1例、及び本発明の可変ビット・レート制御方式の動作ポイントを示す図である。

【図4】本発明の原理に従い、ピクチャのグループ(GOP)に割当てられるビット数を推定する1例に関連付けられる論理の1実施例である。

【図5】本発明の可変ビット・レート制御方法の1実施例のブロック図の1例を示す図である。

【図6】本発明の原理に従い、ピクチャのグループ(G

10

20

30

40

50

OP)に割当てられるビット数を推定する別の例に関連付けられる論理の1実施例である。

【図7】パーC-パーQモデルの1例、及び本発明の別の可変ビット・レート制御方式の動作ポイントを示す図である。

【図8】本発明のメディアン・フィルタリング技術を用いるシーン遷移の間の、GOPビット調整の1例を示す図である。

【図9】本発明の原理に従う、レート制御不安定性のGOPビット調整の1例を示す図である。

【符号の説明】

100 符号器

102 2次元離散コサイン変換(DCT)

*104 動き推定(ME)

106、202 動きベクトル(MV)

108 動き補償(MC)

114 エントロピ符号器

116 エントロピ復号器

118 逆走査

120 逆量子化

122 逆離散コサイン変換

124、502 遅延

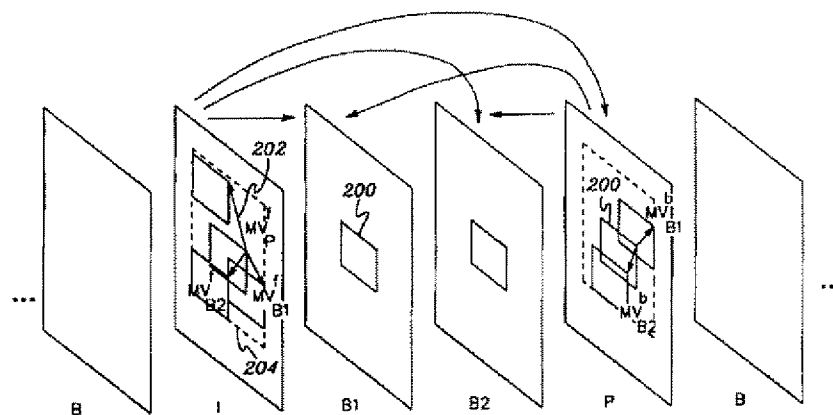
10 200 マクロブロック

204 探索窓

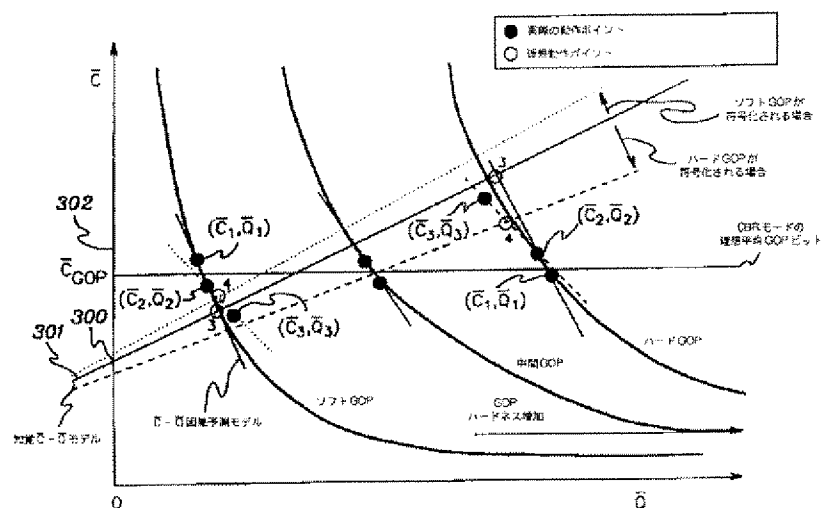
301 知覚ライン

* 500 圧縮エンジン

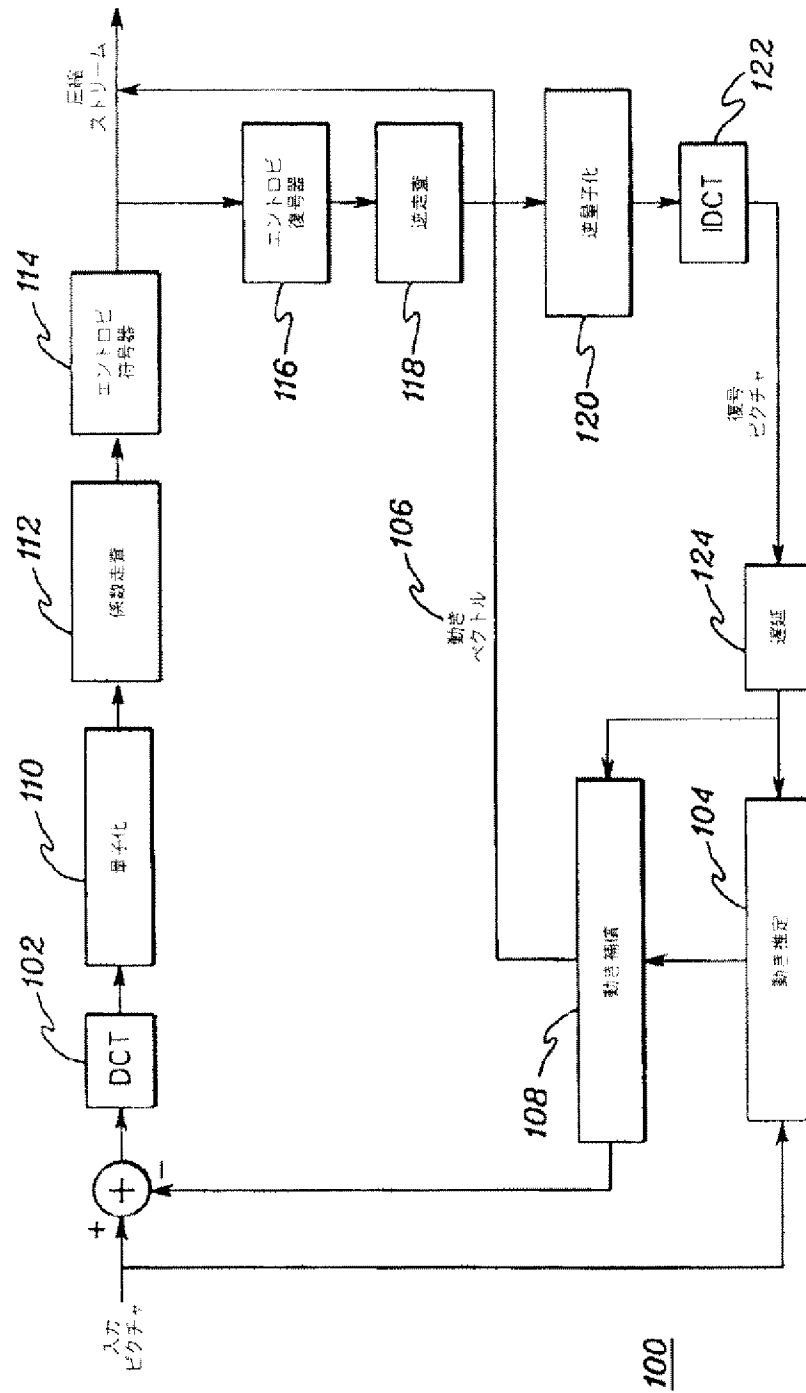
【図2】



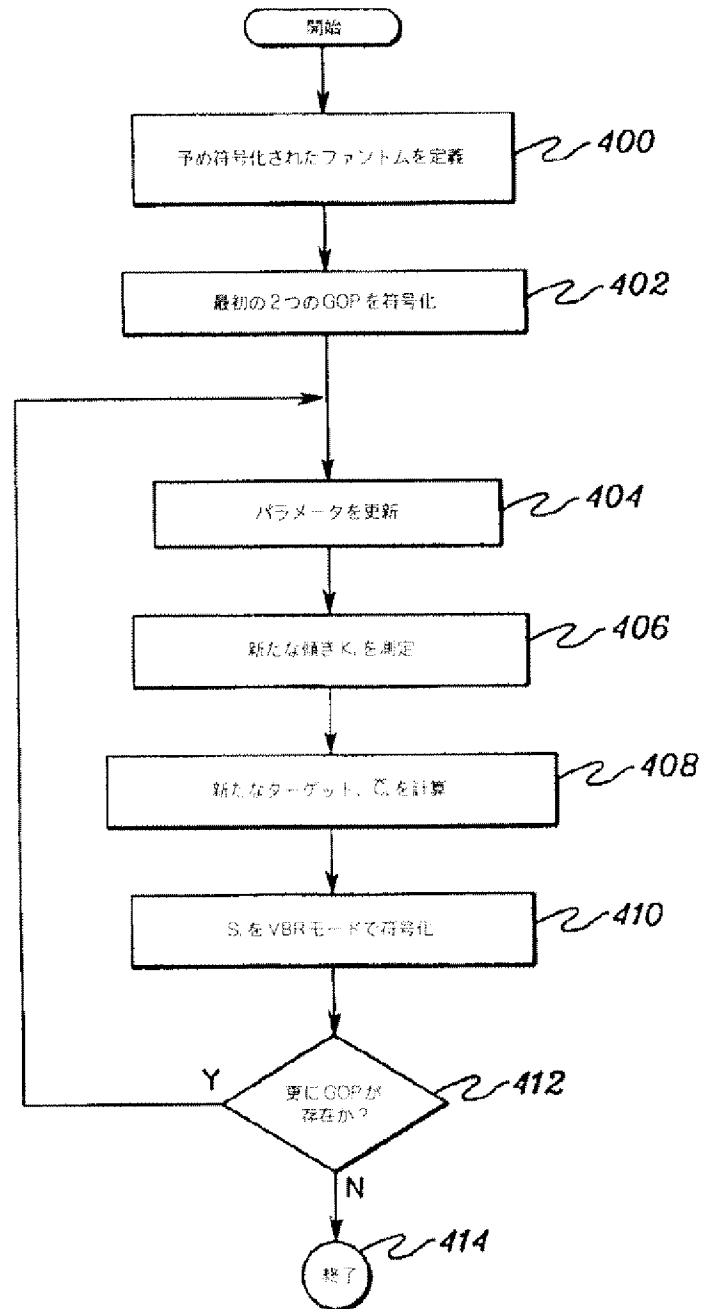
【図3】



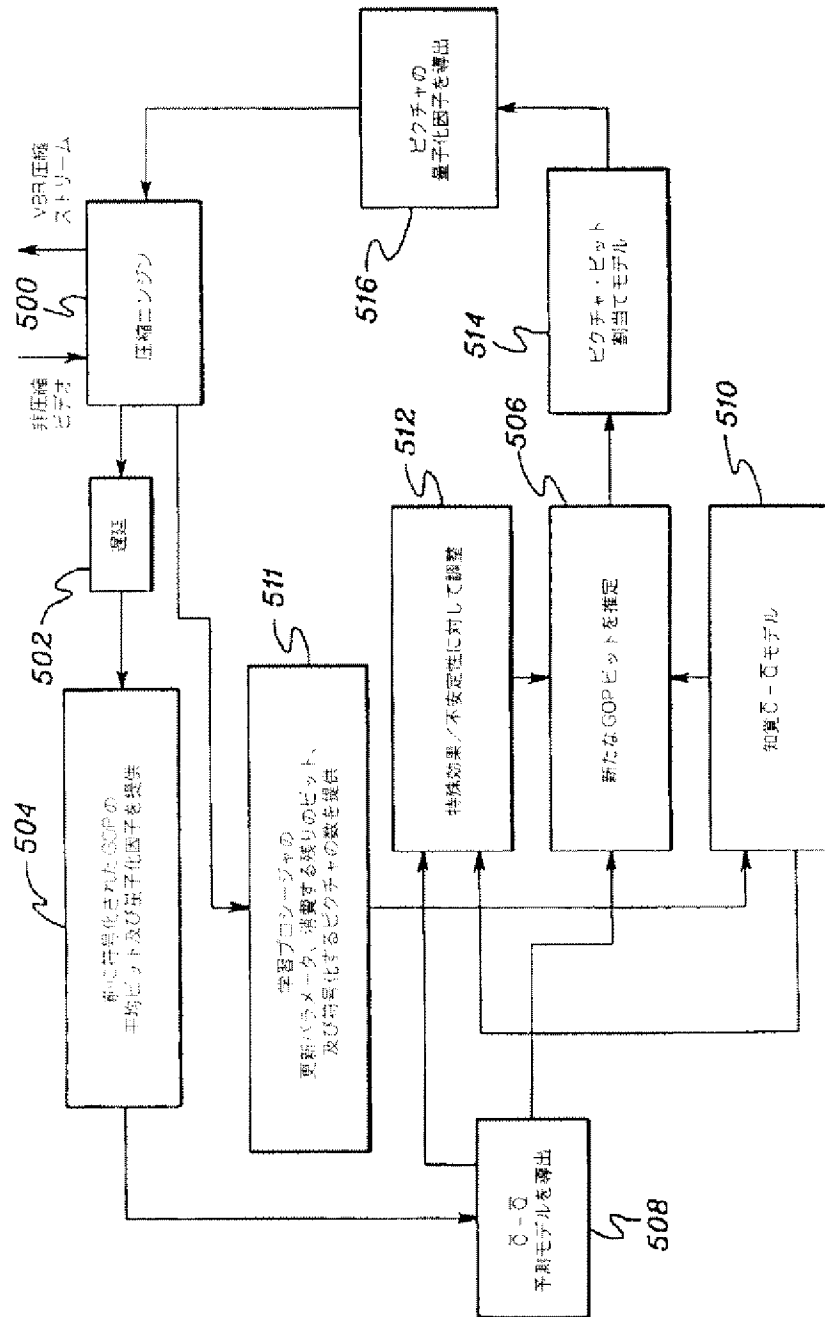
【図1】



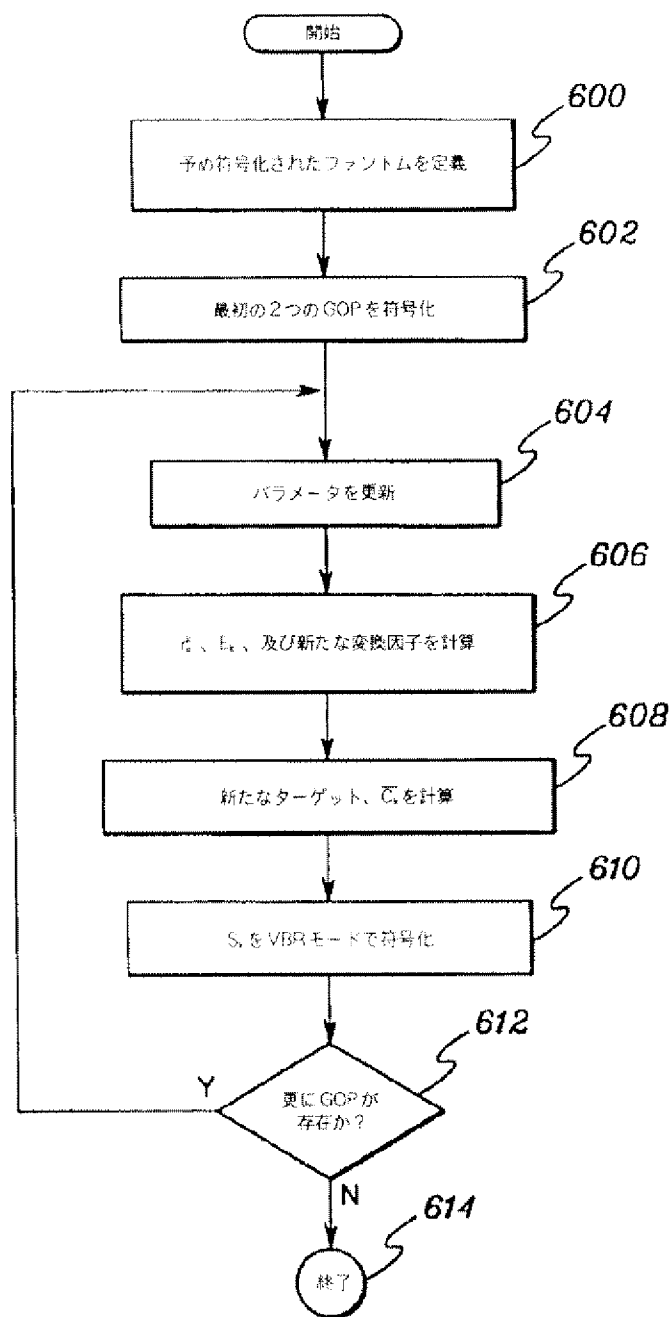
【図4】



【図5】



【図6】



【図9】

